DOCUMENT RESUME

ED 205 174 IR 009 477

AUTHOR	Poncelet, J.
TITLE	Guide for the Establishment and Evaluation of Services for Selective Dissemination of Information.
INSTITUTION	United Nations Educational, Scientific, and Cultural Organization, Paris (France).
REPORT NO	PGI/WS/80/14
PUB DATE	Dec 80
NOTE	78p.: This is a revision of PGI-78/WS/22. Legibility varies.
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Bibliographies: *Computer Programs: *Developing Nations: *Information Dissemination: Information Retrieval: Information Systems: International Programs: *Online Systems: Questionnaires: *Selective Dissemination of Information

ABSTRACT

This guide describes the components of a selective dissemination of information (SDI) service which is designed to give developing countries access to international sources of bibliographic information and provides guidelines for the establishment and evaluation of this type of service. It defines the main features of a computerized documentation system, describes the CDS/ISIS and CAN/SDI software packages which Unesco makes available to member states, and explains the governing principles for system use and evaluation. Lastly, it charts the place of SDI services in the development of a national documentation infrastructure with emphasis on developing nations. Appendices include the results of the evaluation of SDI services in India and Argentina, developed with Unesco's assistance, and the evaluative questionnaires used. Also provided are 13 references specific to the guide and a subject bibliography of 57 references to the literature of information retrieval, networks, evaluation of computerized information services, selective dissemination of information, and statistics. (Author/RAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED205174

UNITED NATIONS EDUCATIONAL,
SCIENTIFIC AND CULTURAL ORGANIZATION

GUIDE FOR THE ESTABLISHMENT AND EVALUATION OF SERVICES FOR SELECTIVE DISSEMINATION OF INFORMATION

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Susanne Adlung

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR009477

(PGI/WS/80/14)
Revision of PGI-78/WS/22

Reproduced by permission
of Unesco

SUMMARY

This guide describes the components of a selective dissemination of information (SDI) service and explains how to establish and evaluate this type of service. It defines the main features of a computerized documentation system, describes the CDS/ISIS and CAN/SDI software packages which Unesco makes available to its Member States, and explains the principles governing the establishment and evaluation of SDI services. Lastly, the guide attempts to chart the place of SDI services in the development of national documentation infrastructures, mainly with a view to improving access for the developing countries to world sources of information. The results of the evaluation of two SDI services developed with Unesco's assistance and the questionnaires used in obtaining those results appear as annexes.

PREFACE

This document is concerned with the problems of computerized documentation, and in particular with the selective dissemination of information (SDI). It deals more specifically with the establishment and evaluation of SDI services designed to give the developing countries access to international sources of bibliographic information. It also takes into account the experience gained in using the CAN/SDI software package in three pilot projects sponsored by Unesco. These pilot projects have yielded information that will probably be particularly useful to anyone weighing the desirability of setting up SDI locally, for they have prompted Unesco, with the assistance of the National Research Council of Canada, to make the CAN/SDI software package available to Member States.

However, in view of the advantages now displayed by international information systems that can be interrogated conversationally, and particularly by bibliographic data bases open to retrospective search, some sections of this guide have been framed to give a clearer picture of the various possibilities offered by those systems compared with SDI services established at the local or national level.

Mr. J. Poncelet of the Documentation Centre of the Institut Français du Pétrole (IFP) prepared this study under contracts awarded by Unesco to IFP. The document was put into final form - with the assistance, in particular, of Mrs. B. Caillaud - in the light of the comments made by a number of experts. We wish especially to thank Mr. R. Gietz, Director of the Centro Argentino de Información Científica y Tecnológica; Mr. J.L. Moya and Mr. E. Molino of the Consejo Nacional de Ciencia y Tecnología of Mexico; Mr. A.S. Raizada and Mr. R. Satyanarayana of the Indian National Scientific Documentation Centre; Mr. E.V. Smith, Director of the Canada Institute for Scientific and Technical Information; and Professor F.W. Lancaster of the Graduate School of Library Science of the University of Illinois, Urbana-Champaign, United States of America.

As its title indicates, this document is more a guide than a set of strict instructions. It represents a synthesis of various opinions which are not necessarily shared by Unesco and which do not commit the Organization. Any comments and suggestions for improving the text are welcome and may be addressed to:

Division of the General Information Programme
Unesco
7 place de Fontenoy
75700 Paris
France.

CONTENTS

		Page
1.	GENERAL CONSIDERATIONS	1
	1.1 Bibliographic information systems: main functions	1
	1.2 Computerization of bibliographic information retrieval	2
	1.3 Problems of access to information	6
2.	UNESCO SOFTWARE: CDS/ISIS - CAN/SDI	7
	2.1 The CAN/SDI component	7
	2.2 The CDS/ISIS component	9
	2.3 Obtaining and installing CDS/ISIS - CAN/SDI	10
3.	ESTABLISHMENT OF A SDI SERVICE	11
	3.1 Coverage - users' and data bases	12
	3.2 Installing the software	14
	3.3 Personnel	14
	3.4 Profile design	15
	3.5 Relations with users	16
	3.6 "Local outposts" of the service	18
	3.7 SDI financing policy	18
4.	EVALUATION OF SDI	19
	4.1 Questionnaires	19
	4.2 Processing and presentation of results	22
5.	THE PLACE OF SDI IN CURRENT OR FUTURE COMPUTERIZED SERVICES	25
	5.1 Establishment of SDI on a local basis	25
	5.2 Use of international services	26
	5.3 Conclusion	27

Page

ANNEX 1

MAIN RESULTS ACHIEVED USING THE CAN/SDI SOFTWARE PACKAGE IN ARGENTINA AND INDIA 29

ANNEX 2

QUESTIONNAIRES USED FOR THE EVALUATION OF EXPERIMENTS WITH SDI IN ARGENTINA, INDIA AND NEIGHBOURING COUNTRIES 47

REFERENCES 67

BIBLIOGRAPHY 68

1. GENERAL CONSIDERATIONS

Technical, economic and scientific activities generate a large stock of information. This may become a prime asset and a deciding factor in a country's development provided that it is processed so as to be made accessible. Only organized information systems can cope with a constantly growing mass of information and effectively perform the process of communicating information. (1)

Information is transferred in a cycle made up of several different stages: the work of the author; publication of the (primary) documents; acquisition of the primary documents by information centres; and their organization and control by those centres (cataloguing, classification, indexing and abstracting and establishment of data bases or banks). These different operations make it possible to disseminate secondary information and to supply primary documents to users on request.

Secondary information services do not transmit original texts of information (primary documents) but bibliographic information which facilitates the retrieval of the information contained in those texts: in other words, they transmit information about information. This secondary information and the services producing it, especially selective dissemination of information (SDI) services, form the subject of this study.

Access to the original documents, which should be foreseen for any secondary service, entails setting up additional structures- document collections, interlibrary loan services and so on- but this aspect of the question cannot be examined in detail here.

1.1 Bibliographic information systems: main functions

The functions of bibliographic information systems are described in very general terms, for in practice a large number of variants are often encountered in adapting the systems to users' requirements and to the constraints of the equipment and methods used. The main functions of information systems follow one another in time in an established sequence, each function constituting a stage that conditions the next.

The acquisition or collection of information. This is the process of selecting the data to be recorded.

Processing. This function involves preparing the bibliographic description of the document (author, title, publisher, date of publication, language, etc.) and describing the document's contents (abstracting or indexing). These data (bibliographic descriptions, indexing keywords and abstracts) are stored and will be used when a search is being made. The purpose of indexing is to extract the information content of the documents by means of concepts that are representative of that content.

The choice of indexing method (free or controlled vocabulary, choice of documentary language) is extremely important and affects the future working of the system. Most information retrieval systems, whether completely manual, semi-automatic (with edge-notched or punched cards) or fully computerized, involve analysis of the documents by keywords. Faulty analysis, which may be due to indexer error or to defects in the documentary language used, will produce a faulty reply when a search is made.

Retrieval and dissemination. Search operations are conducted on the data contained in a catalogue, an index, a secondary document or a computer memory: bibliographic data, abstracts and keywords. Information systems are never specific enough to give ideal answers to the questions asked. Two factors may give rise to faulty answers: noise (the documents identified do not answer the question) and

silence (relevant documents forming part of the document base have not been retrieved). Although this guide is based largely on experience gained with computerized documentation services, it is important to remember that good retrieval and dissemination services can also be established without using a computer.(13)

A search is described as retrospective when all or part of the documentation compiled in the past on a given problem is interrogated. A retrospective search may be personal (when the applicant conducts the search himself) or delegated (when the search is entrusted to an information specialist).

Another type of retrieval function is known as current awareness. In this case, the purpose is to obtain information on new developments in a particular discipline. This function is sometimes described as one of information dissemination. Dissemination may be immediate (in the form of a display) or systematic (in the form of descriptive or abstract bulletins, selective dissemination by group profile or selective dissemination by individual profile).

1.2 Computerization of bibliographic information retrieval

Computer techniques have found significant applications in the documentation field. Although the computer can be used to perform almost all documentary functions, such as managing the orders and the acquisitions file and performing processing operations such as analysis and indexing, its main field of application lies in the retrieval and dissemination of information. The computer can produce indexes that are both standard products and research tools. It can manage data bases and data banks. For retrieval purposes, it can perform retrospective searches of information on line (usually in conversational mode) or in batch mode. Lastly, it facilitates current awareness services by issuing descriptive bulletins and, more important, by providing users with a selective dissemination of information (SDI) service.

The ever-growing mass of documentation is making it essential to turn to the computer for two reasons. Firstly, manual systems become unusable when the mass of documentation reaches a certain size; and secondly, the larger the system, the more accurate it has to be, and it must therefore make use of a larger number of keywords to express the content of the documents and to formulate questions.

Data storage - data bases and data banks

A computerized documentation system is made up of hardware, software (the set of instructions which the computer carries out in performing its tasks), and the data on documents, which may comprise millions of references. As we have seen, these data, stored in the computer memory and retrieved at the user's request, are of two kinds: bibliographic data proper, from which every document can be identified, and information on the documents' contents. The data are "fed" into the computer after being entered on a schedule and coded. They then form a file or a set of files (often called a data base) which can be identified by discipline and by origin. In this paper, the bibliographic data bases created by the accumulation of bibliographic records are merely called data bases and contain secondary information. A data bank is the result of computerization of one or more data bases which may originate from different sources. In the documentation field, the terms data "base" and data "bank" are often used indiscriminately in practice.

The documentary language often forms an integral part of a data base, constituting the keyword file or the thesaurus file. If so, the list of keywords - of which there may be thousands - or the thesaurus is recorded in the computer memory. It requires frequent updating by the addition or deletion of terms and the improvement of inter-term relations.

Document files may be of two distinct types:

- the sequential file, in which the documents, with their keywords, are recorded one after another in order of acquisition. This system does not allow direct access and entails reading the entire file in order to retrieve the information from it, so that the search takes longer than when an inverted file is used;

- the inverted file, organized by fixed order of keywords, each keyword being followed by the identification numbers of the documents indexed by that keyword, thus enabling the computer to retrieve the relevant documents direct. This system is accordingly quicker, but the high cost of setting up a large inverted file is only warranted by a heavy demand for information.

In most cases the sequential file is the basic file from which an inverted file can be derived. In computerized documentation systems it is possible to produce several inverted files (based on significant words in the title, authors' names, and so on), thus increasing and refining the possibilities of direct-access search.

Interrogation of files - on-line or batch interrogation

The interrogation of a data base in accordance with users' requests entails strict compliance with the rules for writing the question and the logical rules imposed by the system. The interrogation proper makes use of the logical operators of Boolean algebra. Keywords representing concepts are linked by logical operators (those most used in documentation are AND, OR and AND NOT) to produce search equations (see Figure 1). These equations are used by the computer to find the references to presumably relevant documents which are communicated to the inquirer.

In batch mode interrogation, the reply is not given at once. The questions are assembled and coded in advance until a "profitability threshold" is reached, so that the computer can be operated thriftily. It is important to formulate the questions properly, since they cannot be asked again at once in the event of an error, but only on sight of the first answers.

On-line interrogation (usually performed in conversational mode) makes it possible for a request for information to be answered almost immediately; on-line interrogation is faster and more flexible but also more expensive than batch mode interrogation. With conversational systems a dialogue can be conducted with the computer through a terminal, whether the computer is a few metres or thousands of kilometres away.

Most computerized documentation systems are designed to work in both batch mode and conversational mode. However, there is an increasing tendency to work only with "on-line" systems. The number of references expected and considerations of cost and speed determine the best type of interrogation.

Batch processing is cheaper and is justified mainly in cases where individual searches are not urgent, or where they require considerable computer time, or else where it is possible to set up a permanent file of search equations (SDI utilization).

On-line processing is more suitable for urgent and non-repetitive document search processes such as retrospective search, but it requires special hardware and software that have been adapted for the purpose. However, some on-line systems can also offer an SDI service, usually by providing facilities for the permanent storage of profiles that can be recalled by the users at any time.

Figure 1

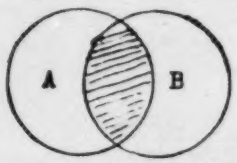
LOGICAL OPERATORS AS A BASIS FOR SEARCH EQUATIONS

Equation 1/

A AND B

AND function (intersection)

The hachured zone represents the set of documents which are:

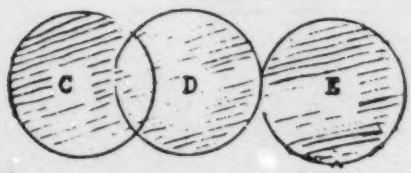


indexed simultaneously by keywords A and B

Equation 2/

C OR D OR E

OR function (union)
The OR is inclusive

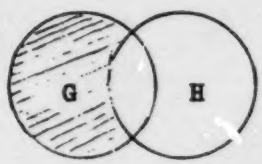


indexed by at least one of keywords C, D and E (any combination of those three keywords)

Equation 3/

G AND NOT H

Exclusion function



indexed by the keyword G to the exclusion of keyword H

N.B. The notations may change, depending on the software being used. The functions, i.e. AND (simultaneous presence), AND (inclusive-OR) and AND NOT (exclusion) remain the same.

Terminals (card readers, teleprinters, visual display terminals and so on) are generally used to transmit the requests to the computer and receive the answers. In long-distance communication (teleprocessing), the information carrier signals (digital signals converted, if necessary, into audio signals) use the switched telephone network, a telegraph line or a local cable, and sometimes pass through a specialized data-transmission network before they finally reach the recipient computer. In such cases the processing centre must have input-output equipment which converts the signals of the request into a form that the computer can handle for search purposes and returns the answer to the inquirer by an appropriate mode of transmission.

Although teleprocessing can be used for requests for information both in batch mode and on-line, its most significant application in the information field - where it has indeed become an international industry - is the on-line interrogation of large data bases through computer communication networks. These networks consist of a set of computers which may be in the same country or, as is increasingly often the case, in different countries. Telecommunication links provide remote access to one or more data banks.

The structure of computerized services

Retrospective search. Of the mass of documents in circulation, some lose their relevance quite soon, while others continue to be useful for more than 10 years. The latter ensure that data bases may be of interest to a user of bibliographic information, such as a research worker, at any time. A user will ask his question during an isolated original search for documents on a given subject. Normally, a documentalist intermediary will reformulate the question in accordance with the procedures of the system used (keywords, logical operators, etc.) for processing in batch mode or on line. The search for the information will be carried out on sequential or inverted bibliographic files, usually held in disc memories and accessible very quickly.

Selective dissemination of information. This type of information service consists of the periodic processing of new information received (current documents) and their dissemination on the basis of profiles. The "profile" concept is fundamental.

Each profile is a (symbolic) description of a documentation requirement. It is represented by a permanent search equation. Since the search affects a limited number of documents (those recorded since the last consignment, i.e. since the previous processing run), the standards of accuracy are less strict than in retrospective search operations and this must be taken into account in formulating the equation. In all cases, the user is regularly supplied with information concerning newly recorded documents corresponding to his profile. In such computerized systems, a file of the profiles is normally established and regularly checked.

In addition, a distinction is usually drawn between two types of profiles:

- the standard profile (or macro-profile) is designed for a group of users (such as a department, working group or administrative unit) who share a common constant need for current documentation over a given period;

- the personalized profile is designed for an individual user whose type of work, discipline and aims impose strict limits on his documentation requirements.

The document references corresponding to the themes are selected, edited and dispatched to their respective users in different forms. Dissemination varies with the type of profile. In the case of the standard profile, the print-out consisting of all the bibliographic references with their keywords (and abstracts if these are included in the data base) is reproduced and sent to the corresponding subscribers.

In the case of personalized profiles, other items may be added to the list of bibliographic references; photocopying the first page of every document in the profile is a widespread practice.

With a national or local data base, the computer procedure for handling data searches for selective dissemination of information is essentially the same as that used in retrospective search, i.e. access to sequential or inverted files by on-line or batch processing, depending on the hardware and software available.

It is also possible to set up a SDI service on the basis of current bibliographic information received from external sources without having to incorporate that information in a national or local data base. Magnetic tapes are the most commonly used medium for this type of service. By using appropriate software, it is possible to compare the contents of these tapes with the themes of interest (profiles). Batch interrogation is normally used and, once the operations of comparing references and profiles have been carried out, the computer is free to do other work.

1.3 Problems of access to information

Improving access to information in a country entails setting up appropriate structures at the local or national level and adapting them to international structures. In the first case, the task is essentially one of managing the information that is inside the country itself; in the second, of making the best possible use of information of general interest chiefly obtained from external sources. The pursuit of these two aims entails making a choice of national strategies for the preparation, sorting and dissemination of information, and such choices present problems of training, adaptation and costs.(1)

Local services

In national terms, the identification and use of national documentation through services such as SDI and retrospective search often have priority. The medium-sized files needed for this type of project can be managed with relatively simple hardware and one of the many documentary software packages that have been tested and are now available, including several versions of the ISIS software (see section 2.2).

Universal services

Universal secondary systems are systems supplying information of interest to several countries or to the international community as a whole within a particular field of knowledge. As to the organization of these systems, the input data are collected by a number of separate national centres or else by a large information centre having users in several countries. These input data are processed to produce a machine-readable data base, which is used in turn to produce abstract bulletins and other international secondary publications, as well as tapes for subsequent use in computerized SDI services and retrospective search.

International systems in one particular category - co-operative systems such as AGRIS (International Information System for the Agricultural Sciences and Technology) or INIS (International Nuclear Information System) - are of special interest to developing countries wishing to gain access to information in certain specialized fields. The data bases of these systems have international or regional coverage and are accessible through co-operative channels.

Some companies and other institutions also make available to users a set of such data bases through remote interrogation. The international data bases are fed into one or more data-processing units called hosts, which are often accessible in conversational mode. These big international hosts are increasingly popular. They are a

practical proposition provided that they have a large number of users, but they are still relatively expensive. Consequently almost all the big hosts are located in the industrialized countries at the present time, and conversational access to them from developing countries is often very expensive or unreliable. Where this is the case, access to such centres by correspondence to meet document search requests may be a worthwhile alternative.

As a rule, the establishment of autonomous data base services is beyond the means of a small country or a single documentation centre even on the basis of subscription to existing data bases on magnetic tape with the aim of setting up universal bibliographic search services. However, for purposes of retrospective search or SDI, use can be made of the traditional indexed bulletins (secondary publications), specialized centres which process and sort documentation, or on-line links with the international hosts.

For SDI alone, the use of magnetic tape resources from large data bases is also possible for a centre which can afford complete management of only small bases in their entirety. Hence the establishment of computerized SDI services has aroused particular interest in the developing countries and Unesco has participated in establishing several national SDI services. This guide has been written to enable other countries to profit by the experience thus gained.

2. UNESCO SOFTWARE: CDS/ISIS - CAN/SDI

When a SDI service is to be established, the acquisition of appropriate software on reasonable terms is often a major difficulty. The development of software in situ is a process requiring not only confirmed expertise in the field of documentary data processing but also a great deal of time and substantial resources. Many SDI software packages are available on a variety of terms from bilateral, international commercial sources. The choice of software depends on a number of factors, and especially on the hardware available or being acquired and on the characteristics of the planned service. Unesco's dissemination of a specific software package, CDS/ISIS - CAN/SDI, is described here as a tangible example of international action in this field and in order to convey an assessment of software problems in the context of this study.

The combined CDS/ISIS - CAN/SDI software package consists of two components capable of performing a number of documentary functions:

- the CAN/SDI component provides effective software for SDI services using universal bibliographic bases on magnetic tape;

- the CDS/ISIS component affords a means of managing local bibliographic files for SDI and retrospective search by either batch or on-line processing. This component can also be used, in theory at least, to integrate universal bibliographic data into local data bases.

2.1 The CAN/SDI component

CAN/SDI is a documentary software package originally developed by the Canada Institute for Scientific and Technical Information (CISTI). It provides for the selective dissemination of information from bibliographic data bases on magnetic tape, the choice of bases being left to the discretion of those responsible for each SDI service. It should be explained that, in its very principle, CAN/SDI software is not designed to perform retrospective searches. It compares the profiles stored with only the latest delivery from the required data base and does not build up information from the base for future searches.

In Canada this software is used for the selective dissemination of information from a large number of data bases. At the time of writing this guide, they included the following:

- AGRICOLA (bibliography on agriculture, formerly called CAIN)
- BIOSIS PREVIEWS (Biological Abstracts and BioResearch Index)
- CHEMICAL ABSTRACTS SEARCH (superseding CA CONDENSATES)
- CHEMICAL TITLES
- COMPENDEX (Engineering Index)
- ERIC (Educational Resources Information Centre)
- FSTA (Food Science and Technology Abstracts)
- GEOREF (Geological Reference File)
- GRA/NTIS (Government Reports Announcements/National Technical Information Service, United States of America)
- INSPEC (Information Service in Physics, Electro-technology and Computers)
- MARC II
- METADEX (Metals Abstracts Index)
- PSYCHOLOGICAL ABSTRACTS
- SCI (Science Citation Index)
- SSCI (Social Sciences Citation Index).

It would be relatively easy to adapt the software to process other bases, such as tapes from the AGRIS and INIS international systems, which conform to the standardized international format for the exchange of bibliographic information (ISO 2709).

In Canada the number of subscribers served by the national SDI service using this software was 2,100 at the end of 1977, compared with 1,800 some three years earlier. On the average, each subscriber uses 1.5 bibliographic bases. In 1975-1976 five million references extracted from these bibliographic bases were printed and sent out to subscribers.

Later on CISTI was willing to make the CAN/SDI software package available outside Canada and offered its technical collaboration to those wishing to install it.(2) By the end of 1978 seven organizations had taken advantage of the offer. These included organizations in three developing countries responsible for national SDI services, which received initial assistance from Unesco under a pilot project:(3)

in Argentina, the Centro Argentino de Información Científica y Tecnológica (CAICYT) of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET);

in India, the Indian Institute of Technology, Madras, in collaboration with the Indian National Documentation Centre (INSDOC);

in Mexico, the Consejo Nacional de Ciencia y Tecnología (CONACYT).

Users in a number of neighbouring countries have been associated with these projects and have provided a number of documentary search themes for processing. Thus Bolivian users took part in the Argentine experiment, and users in Nepal and Sri Lanka were associated with the Indian project.

During the one-year pilot period, the three projects provided a SDI service using the Chemical Abstracts Condensates data base. A detailed evaluation of two of the projects - those of Argentina and India - gave positive results (presented in Annex 1) and at least two of the services - those of India and Mexico - are now proceeding, without Unesco's assistance, to extend their SDI service in terms both of number of users and of number of data bases accessible to interrogation. A fourth pilot project was launched in Nigeria under the auspices of the National Science and Technology Development Agency in 1979.

In view of the success of these initial experiments and of the interest shown by Member States, Unesco and CISTI have agreed to distribute the CAN/SDI software as an integral part of the CDS/ISIS documentary software package already widely disseminated by Unesco.

It should be mentioned that CISTI has published two very detailed manuals on CAN/SDI. The first of these, the CAN/SDI Profile Design Manual,(4) contains the following sections:

a general introduction to the system;

how to design a CAN/SDI profile correctly, explained in 13 stages;

a glossary of the concepts necessary for profile design;

how to improve profiles;

analysis of the problems specific to each data base to which CAN/SDI is adapted.

The second manual, more condensed, deals with the data-processing aspects. This is the CAN/SDI System Manual, which is obtainable from Unesco.

The CAN/SDI software operates on the "OS" operating systems for IBM computers such as those in the 360/370, 3030 and 4300 series, and on the same company's "TSS" system.

2.2 The CDS/ISIS component

The original version of the ISIS software, developed by the International Labour Office during the 1960s,(5) has since been refined considerably in co-operation with several other organizations.(6) It is now being used by a large number of organizations all over the world to manage bibliographic information services. A new version of this software, developed for Unesco's documentation services under the name of CDS/ISIS (Computerized Documentation System/Integrated Set of Information Systems), is offered by Unesco to Member States.

With CDS/ISIS it is possible to create and manage medium-sized data bases, retrieve information and publish a large number of bibliographical products such as indexes and descriptive bulletins. The software operates by batch processing on the "OS" operating system of the IBM 360/370, 3030 and 4300 series computers and on compatible hardware such as the RIAD 1035, 1040 and 1060 computers. It uses the IBM/CICS system for on-line processing. The inverted file procedure is normally used for information retrieval.

The size of each data base is at present limited to 100,000 records (i.e. 100,000 document entries) for which the CDS/ISIS component can perform complete documentary processing, including SDI and retrospective searches. Although the software is designed for managing local data bases, it is also possible to integrate a limited number of references from external sources (such as those received for SDI operation) into the CDS/ISIS files. The number of such references could be increased by modifying the software if management of these universal data bases in their entirety should prove desirable (see section 5).

It should also be mentioned that the CDS/ISIS software can process many kinds of data which are not strictly bibliographical, such as registers of current projects, lists of institutional sources of information and certain non-documentary files concerned with various aspects of administration.(7)

By the end of 1979 CDS/ISIS had been installed for a variety of applications in 24 national institutions and four international organizations. Complete documentation on the system is available. It includes a general description(8) which includes the terms offered to organizations wishing to acquire the software.

2.3 Obtaining and installing CDS/ISIS - CAN/SDI

Since 1979 the combined CDS/ISIS - CAN/SDI software has been made available to Member States on the terms formerly applied to the CDS/ISIS software alone.

After agreements have been signed between Unesco and any institution wishing to set up computerized documentation services, Unesco organizes, within the limits of the places available, international sessions of training in the use of the CDS/ISIS or CAN/SDI software or of both components of the CDS/ISIS - CAN/SDI package. The Organization also makes the necessary arrangements for transport and installation.

If universal data bases are to be processed with the CAN/SDI component, it is essential to provide for additional introductory training in profile design to suit the data base or bases that the service is to use (see section 3). This type of training has to be arranged at appropriate SDI centres, as Unesco has no experience in that field.

No direct financial charge is made for the CDS/ISIS - CAN/SDI software. However, the travel and subsistence expenses of the people given training in the use of documentary systems will continue, in principle, to be borne by the organizations applying for one of the systems. In the case of developing countries, Unesco will endeavour to find financial assistance for the training courses either out of the funds provided for in its own programme or else from extra-budgetary sources, and will also try to facilitate access to the international bibliographical bases and to training for profile designers.

Once a software package has been installed, the applicant organization is provided with technical assistance, which may take one or more of the following forms:

visits by Unesco experts if necessary;

technical advice by mail, telex or telephone;

regular deliveries of software updatings and related documentation.

The three basic requirements for participation in this programme are as follows:

1. an analysis of the requirements for documentation services, prepared from a survey of the groups of users to be served;

2. the infrastructure needed to run the services properly, i.e. appropriate documentation facilities and staff;

3. access to suitable hardware.

Before a decision is taken to acquire CDS/ISIS - CAN/SDI software, it would be useful to make preliminary contact with organizations having already installed the component or components envisaged. These preliminary contacts may be confined to exchanges of correspondence or requests for practical advice, but they may also extend to spending an appropriate length of time on the spot. Any request for information on the CDS/ISIS - CAN/SDI system may be addressed to the Division of the Unesco Library, Archives and Documentation Services, 7 place de Fontenoy, 75700 Paris, France.

3. ESTABLISHMENT OF A SDI SERVICE

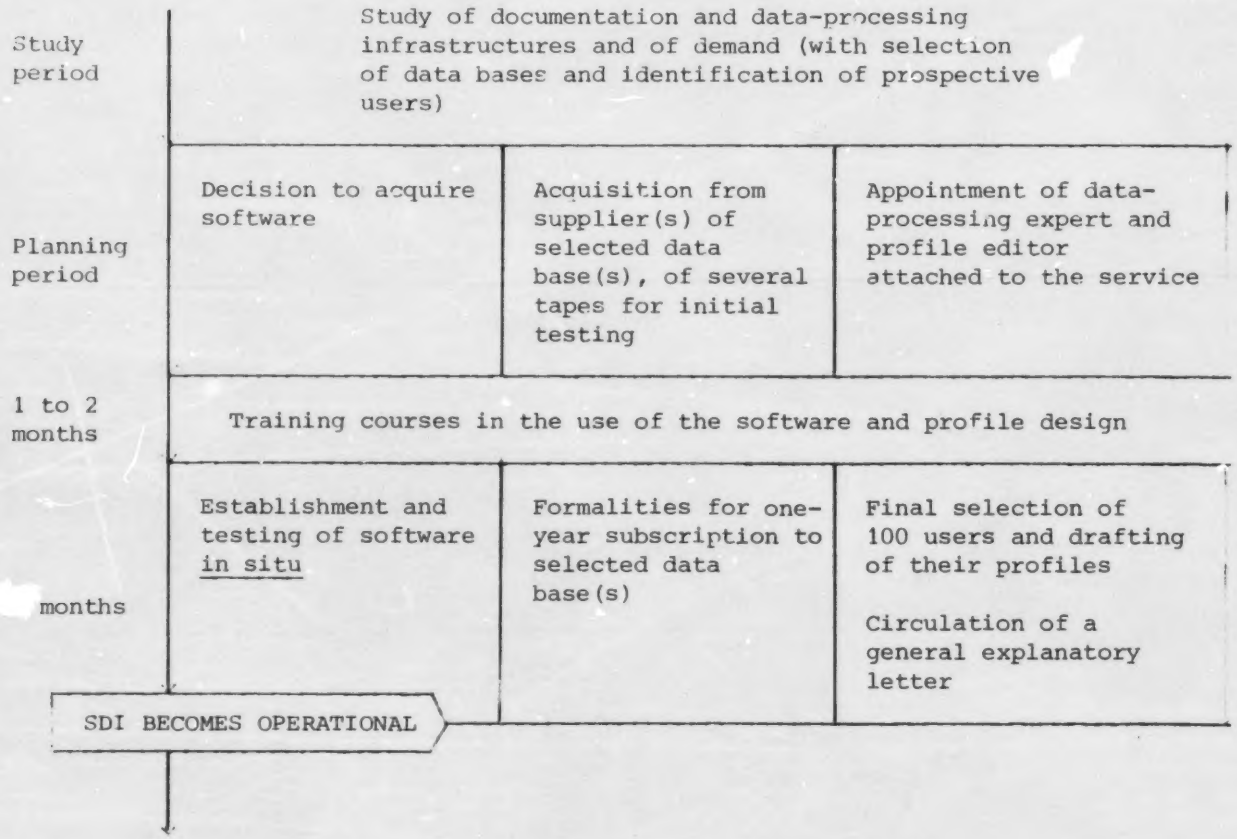
The establishment of a SDI service requires careful preparation, considerable expertise in documentation and data-processing, and appropriate infrastructures.

The establishment of a SDI service operating on a universal data base formed of periodic magnetic tapes will be discussed in the following pages. The establishment of a SDI service operating on local data bases gives rise to fewer problems as regards access to information and scanning techniques. By contrast, the problems involved in setting up and managing local data bases are very real and have to be given close attention.

Figure 2 shows, in diagrammatic form, the preparatory stages in the establishment of the SDI service. These stages will be discussed below, together with several other aspects of the development of an operational service.

Figure 2

DIAGRAM OF THE PREPARATORY STAGES
IN THE ESTABLISHMENT OF A SDI SERVICE


Study period
Study of documentation and data-processing infrastructures and of demand (with selection of data bases and identification of prospective users)
Planning period
Decision to acquire software
Acquisition from supplier(s) of selected data base(s), of several tapes for initial testing
Appointment of data-processing expert and profile editor attached to the service
1 to 2 months
Training courses in the use of the software and profile design
months
Establishment and testing of software in situ
Formalities for one-year subscription to selected data base(s)
Final selection of 100 users and drafting of their profiles
Circulation of a general explanatory letter
SDI BECOMES OPERATIONAL

About one hundred users can be regarded as a reasonable basis for starting up a SDI service. It is advisable to maintain that number during an experimental stage lasting some 11 to 12 months, which is the length of time needed to run in and evaluate the service (see section 4). For the same kind of reasons, it is also advisable to confine operations to a single data base during the experimental stage, especially if the SDI service represents the first venture into computerized documentation.

The final goal is to have several hundred profiles and perhaps even to extend the area of coverage of the service, and this may require several years' prospection, depending on circumstances. Generally speaking, the higher the annual cost of subscribing to the magnetic tapes, the more advisable it is to try to spread the cost over a large number of profiles. (The annual cost of subscribing to a data base may be in the US $5,000-10,000 range in instances unless preferential rates are granted or the data bases, as in the case of AGRIS and INIS, are made available to certain official bodies free of charge.) The possibility of extending the services to users in neighbouring countries, or indeed to the region, may appear desirable in the long run.

3.1 Coverage: users' and data bases

The subject or subjects likely to be covered by a SDI service will be defined in terms of national priorities, main user categories and the terms of reference of the centre organizing the service. Users of this type of service will, as a rule, be researchers or persons responsible for management or research questions, but other groups such as teachers, technical personnel and heads of government departments may also be interested.

It is most important in the initial planning stage to undertake a preliminary study of the existing infrastructures capable of providing back-up for the SDI service (such as available document collections in the particular area selected, computers, specialized staff, and so on) and of potential demand for a service of this type (i.e. the likely number of users in the subject-area or areas envisaged). At this stage, the various bibliographic data bases capable of being used will have been identified. If the infrastructures turn out to be adequate for the establishment of a SDI service, the study should make it possible to make a decision concerning:

the community of users to be catered for in the first instance;

the data base to be used.

It is not possible here to evaluate how long the study will take, since this depends on local conditions, the information sources available, and so on.(9)

Once the decision to establish a SDI service has been taken, it is important to find the first users as quickly as possible. The different sectors should be canvassed and advantage can sometimes be taken of the contacts established during the initial study. If no such contacts have been established, canvassing is liable to take much longer. Making the existence of a new service known to the largest possible number of potential users is an undertaking whose difficulty should not be overlooked. As a rule, a prospectus is sent out to organizations, associations and private individuals likely to be interested in such a service, but other forms of publicity can also be envisaged.

It is possible to ensure that a satisfactory match will be struck between search themes and the bibliographic data base by making an appropriate selection of users in the first instance. In other words, a prior attempt will be made to seek users whose search themes will be well covered by the prospective bibliographic data base, the aim of this being to increase the chances for success of the experimental stage, which will entail considerable difficulties in any event. An approach along these lines obviously does not mean that it is wished to deprive some users of the benefits of SDI; the idea is merely not to prevail upon people to take out a subscription which they will certainly not wish to renew. It should be stressed, however, that when such an initial selection is made, it cannot be claimed that a representative sample of the total user population is being chosen. This lack of representativity would, to some extent, be liable to distort the forecasts on subsequent development of the service made on the basis of the pilot experiment.

Two cases can be envisaged in the course of the selection process:

- the proposed base is already available to potential users as a secondary review in paper form;

- the base is not available locally in that form.

In the first case, users will normally scan the secondary periodical, and they can be asked to indicate to what degree the periodical gives satisfaction as far as their own search theme is concerned, in accordance with the following table:

	From the standpoint of the quality of articles	From the standpoint of the number of articles
5	periodical excellent	periodical excellent
4		
3	periodical average	periodical average
2		
1	periodical poor	periodical poor

When the users are being selected, the choice will go to those having given the most favourable answers (as measured, for instance, by the product of the two marks for quality and number).

In the second case, when the data base is not available on paper, the person responsible for the survey can:

- either describe the contents of the base being offered to potential users as exhaustively as possible and then ask them for their assessment (on the above 5-to-1 scale). The risk is that they will be unable to give an accurate answer;

- or else scrutinize very closely himself the search themes provided by interested users, although this calls for an excellent knowledge of the data base.

This test can also be applied to several bibliographic data bases when the initial study of the potential demand is being made. It will then make it possible to choose a user group/data base combination that is suited to the experimental phase. Other variants of this method can also be devised.

It is, of course, quite possible to maintain a completely open-minded attitude and to accept any search theme coming within the scope of the service without selection. This approach has the advantage of reducing the initial work-load and of making it easier to draw conclusions from the experimental phase with an eye to the future development of the service (reduction of possible distortions due to the user sample). If there has been no selection process, it will be preferable to accept a higher number of users - say 125 to 150 - during the experimental phase, in anticipation of drop-outs.

The formalities involved in taking out a one-year subscription to the chosen data base may be rather long and complex. It is advisable, therefore, not only to reach agreement with the supplier, but also to make proper arrangements to ensure delivery of the tapes, including any customs formalities required. The diplomatic pouch could perhaps be used as a temporary expedient, in which case all the people concerned should be given precise instructions, failing which the tapes may be "held up" somewhere because nobody has been given written instructions or because those instructions have not been passed on.

3.2 Installing the software

The probelem of selecting or developing software for SDI was touched on briefly in section 2 above. The choice will have to be made in the light of the local situation in every instance. It is important however to ensure that the software is indeed available and that it will have been tested in good time.

The experience of Unesco pilot projects demonstrates that allowance has to be made for a period of several months for installing and thoroughly testing the software, even in cases where this has been "run in" at some other SDI centre that is already operational. It must also be borne in mind that it is essential to have a complete set of documentation available and to have completed a basic training course beforehand.

As a general rule, it is not necessary to take out a full subscription to the data base when the software is being tested on first being installed. It is sufficient to have a few magnetic tapes, which may even be old ones, provided they have the same format as the most recent tapes. It is uaually possible to obtain such tapes from the relevant supplier, by explaining that it is wished to launch an experimental SDI service and that the software has to be tested.

If the data processing is carried out on a computer which does not belong to the documentation centre responsible for organizing the SDI service, as is often the case, problems of coordination are bound to arise. Right from the outset, therefore, it will be very important to formally establish a precise and detailed division of work and to ensure close liason between the units involved. The acquisition of a minicomputer for SDI and other types of documentary applications offers an alternative that can be examined in some cases.

3.3 Personnel

A SDI service must necessarily have the following personnel:

- a data processing expert responsible for the software and its operation;
- a documentalist responsible, in particular, for designing the profiles.

The data-processing expert must have considerable experience in operating the proposed hardware and in programming in the data-processing language or languages of the software. He will be responsible for setting up the software, operating it on a regular basis, servicing it and making any changes necessary. He may be engaged part time on this work, provided his responsibilities are clearly defined. Undernormal circumstances, he should undergo about one month's training in a centre already using the software, during which time he can also familiarize himself with the principles governing SDI. The length of the training period can be reduced if the trainee is already very experienced and if really exhaustive documentation on the software is available.

The documentalist will be responsible for designing user profiles and for maintaining liaison with users on all matters pertaining to the SDI service. He may also perform other duties in the documentation centre responsible for organizing the service, but it has to be borne in mind that the initial user selection process and the design of their profiles will entail several months' full-time work. The post requires a sound background of experience in documentation, especially in traditional bibliographic searching. It is essential for the general standard of education and training of the candidate for this post to be such as to enable him to converse with users as an equal. Furthermore, a training period of about one month in a SDI service using the relevant software and data base will certainly be necessary.

With the passage of time, the service will develo and the number of users and areas covered will keep on growing. This may make it necessary:

- to consult on profile design with outside experts on a part-time basis, such as professors and lecturers from nearby universities, who will be recruited on terms to be negotiated in each individual case;

- to recruit a specialist in the subject-area involved, if warranted by the volume of work;

- to set up "local outposts" for profile design (see Subsection 3.6)

If none of these courses is feasible, each search theme should be described in particularly clear detail right from the outset, in which case it will be for the user to suggest to the SDI centre any changes which need to be made in his profile, which will require more active collaboration on his part than in the above-mentioned situations.

The actual work done by the documentalist, such as drafting schedules, punching cards and checking, is of a conventional nature. It calls for quiet surroundings and a normal work-space. Mistakes are bound to occur when the work performed has to be done too quickly. It then has to be corrected, and this entails tiresome re-checking. Momentary changes in activity enable the personnel involved to regain their ability to concentrate, as is well known.

3.4 Profile design

A list of future users' search themes has to be drawn up before their profiles are designed. The search theme and the name of the organization responsible for it are obviously confidential items of information. The ethics of documentation work require that they should be properly safeguarded and that only the personnel of the documentation service needing them should be allowed access to them/

When profiles are being designed, special account must be taken of the particular features of the data base, such as its structure, indexing language, and so on. However, a number of guidelines can be laid down here.

When describing their fields of interest or search themes, users should first provide the following information:

(a) the actual search theme and its general context;

(b) a list of some five to ten reference titles relating to their search theme;

(c) possible, a list of the expressions and abbreviations generally accepted in their branch of activity;

(d) a list of the type of data which they do not wish to receive (i.e. data in areas that may partly impinge on their search theme)

When the relevant profile is being drafted, it may be useful to proceed in two stages, as follows:

- draw up an initial profile in which no account is taken of unwanted concepts;
- subsequently, when the profile is being revised, take steps to eliminate unwanted concepts which have been reported by the user and which have actually been shown to interfere.

The advantage of this approach is that it will help keep the number of search terms per profile to a minimum, by introducing concepts for deletion only to the extent that they really impair the overall quality of the references retrieved. In fact, users may often provide an excessively long or detailed list of unwanted concepts.

It may be asked why the number of terms per profile should be kept to a minimum. Experience has shown that the product of D x T is of considerable importance in the cost of SDI. D, the number of documents per magnetic tape, is a given quantity and only T, the total number of search terms, is a value that remains, to some extent, under the control of the SDI centre. An endeavour must be made to reduce the value of the term T to a minimum while ensuring a high standard of service.

It should be mentioned that, in some data bases, the information is subdivided into numbered sections, and it may be more economical to indicate the section or sections required (or to exclude those sections generating unwanted references) rather than enumerate the full list of unwanted concepts provided by the user. The value of the above-mentioned term D is effectively reduced in that way.

During the experimental phase of the service, the processing of the results can be greatly simplified by abiding, as far as possible, by the principle of "one user - one profile". If that rule is not applied, there will be two adverse consequences:

- there will be a risk of repeating identical keywords in several different profiles for the same user;

- the task of evaluating the service will become more complicated and the findings less clear-cut (see subsection 4.2).

Even so, cases may arise where exceptions can be made to this rule. For example, if the computer print-outs of a profile entail considerable sorting work for users, it may be preferable to subdivide the profile in cases where it is too general. It has to be borne in mind that sorting done by users also costs money.

Experience has shown that a period of at least four months will be needed to design a hundred or so initial profiles. Some search themes, in fact, may not have been adequately explained by users at the outset and this may have the effect of further extending the time needed. Moreover, all the profiles have to be transcribed on punched cards, which then have to be duly checked. Since the work involved can be done properly only if sufficient time is allowed, it is advisable, when forecasts are being made, to overestimate the time needed rather than underestimating it.

3.5 Relations with users

SDI will be a success only if the confidence of the users has been won and if they are offered a high standard of service. It is particularly important, therefore, to carefully plan and maintain the frequency at which profile print-outs will be issued and to ensure that the profiles are carefully managed. The latter requirement entails a two-way exchange of information between users and the documentation centre before the profiles are first drafted, and that exchange should continue, so as to enable defective profiles to be improved.

When the SDI service starts operating, it will be useful to send out an information circular to all users, with a view to eliciting their collaboration. This circular should explain, in concise and clear terms, what SDI is and what users can expect of it. In addition, it should invite them to send in their comments to the documentation centre as often as necessary, since such comments are the only items of information enabling the profiles, and hence the quality of the service, to be

improved. The circular could be appended to the first questionnaire for evaluating the service, if such questionnaires are planned (see section 4).

Among the efforts to keep users informed, particular importance should be attached to certain features of SDI that are liable to be misunderstood. These include:

(1) The usefulness of the SDI service

A user's knowledge of the state of the art in a given theme may be:

- in arrears of the state of knowledge in the international research community;
- abreast of the main research efforts;
- ahead of the state of knowledge in the international community.

In the first case, before users can really profit by SDI, they will need an up-to-date bibliography on their particulat theme. If the gap that has to be made up is not too large, international retrospective retrieval systems which supply - simplifying matters somewhat - references starting from 1970, make it possible to update the relevant themes very quickly.

If the theme and the user's state of the knowledge are abreast of the international research community, SDI will normally supply almost all of the signifigant references.

Thirdly, if the search theme is ahead of the international situation, the SDI profile will act as what might be termed a "competitor detector", in which case the correct answer will be that there are no patents or references.

(2) The relevance of information from the SDI service

Users should be warned, right from the outset, that it is only very seldom possible to obtain the following results simultaneously:

- 100 percent "precision", i.e. a ratio of the number of signifigant references provided to the total number of references provided equal to 1;
- 100 percent "recall", i.e. a ratio of the number of signifigant computer-retrieved references to the number of signifigant references existing in the data base equal to 1.

Probably the most rational criterion consists in securing documentary search results with 100 percent recall and in tolerating a margin of error in respect of precision, for which a value of 60 percent can be regarded as being very satisfactory. This will accordingly preclude users from having to use manual search methods to go back over the work which the computer has already performed, but in an unsatisfactory manner.

The literature in the field of documentation is curiously silent on this point. In our opinion, users would be rendered a poor service if the goal were 100 percent precision at the expense of recall, for they would come to realize, sooner or later, that the system does not provide them with certain documents which they will have learnt about from other sources. It is better to encourage users to check primarily the exhaustivity of the information provided and to show them that is in their own interest.

3.6 "Local outposts" of the service

It will very probably be found, as experimental SDI services are being established, that some people will intimate their particular interest in the work of designing and improving profiles either to the SDI centre or through comments made in the course of an evaluation exercise (see section 4). This is a regular occurrence and it is usually found that the people concerned work in quite large research organizations with their own documentation centres.

If there are no unsolicited applications of this kind during the experimental establishment of the SDI service, the organizing centre can, in due course, carry out a survey with a view to eliciting such applications.

The main advantage of this lies in the fact that such users wishing to specialize in SDI can enter directly into contact with their researcher colleagues and have detailed discussions with them on profiles. This approach is normally more rewarding than exchanges of correspondence. By being immersed in the working atmosphere of the research organization of which they are a part, the users in question know their themes of interest in every detail. They can therefore provide an effective liaison between their organization and the centre in charge of setting up the SDI service. Part of the latter's work will be done by the local outpost, and the service's workload will be lightened accordingly.

In order to prepare for setting up what might be termed "local outposts", i.e. local correspondents wishing to specialize in SDI, the organizing centre should:

- draft a simple and instructive profile design manual. If appropriate, the examples given by way of illustration would be more forceful if they were relevant to national priorities;

- organize its own profile design training courses for local correspondents. Normally the most suitable time for training collaborators in profile design is from the beginning of the second year following the establishment of the SDI service. Even so, the centre should make preparations for it well ahead of time.

It should be noted that the beneficial effects of local outposts will be felt as their number - and hence the work involved in profile design - increases. However, not all users will always be able to benefit from the assistance of local collaborators, and a large proportion of the requests for profiles will continue to go directly to the organizing centre.

Lastly, the design of a portion of the profiles should be contracted out only to qualified people who have all the requisite skills and who, by virtue of the profiles they design, show that they have properly assimilated the training they have received.

3.7 SDI financing policy

During the experimental phase of the SDI service, it is only natural not to expect users to contribute to the cost of the service. Following that phase, it may be worth while, for obvious reasons, to recover at least some of the operating costs. Contributions required of users can be adjusted according to their status and resources (private industry, government institutions, universities, and so on).

It may also emerge from an analysis of the situation that an operation will be financially profitable only from the time when it has a minimum number of users; in fact, the expansion of a SDI service takes time. There is nothing to prevent the people in charge of the SDI service from trying to secure the partial but continuing financial support of large professional bodies, such as federations of food-processing

industries or chemical industries or else medical research foundations. This financial support could, if necessary, be supplemented by direct or indirect government support during the establishment phase, and this could be maintained until the service is fully developed. It is also possible to envisage providing a SDI service through a government body, since some of the advantages of the service are only completely meaningful and worth while in a national context, particularly in terms of the national balance of payments in sectors dealing with information and with exportable processes.

Canada is one example of a country in which the government provides widespread support for SDI activities through appropriate institutions. A less costly form of government support could, for example, consist in charging very preferential rates per hour of computer-time, if the computer belongs to a government department. In other countries, various methods of indirect support are used, as in cases where users can take out a subscription to a SDI service and are allowed to count the cost involved as being professional expenses, so that their taxes are reduced accordingly.

These various forms of support from different bodies can easily be justified; computerized documentation services enable users to save time and provide them with the information needed for national development. If an appropriate SDI service caters for a sufficient number of users for any significant length of time, it cannot fail to bring its rewards.

4. EVALUATION OF SDI

The establishment of any documentation service, especially at the national level, represents a significant investment in human and financial resources. It is important, therefore, to ascertain what the value of such a service is by carrying out a thorough evaluation. It was felt that it would be useful here to present the methodology developed for Unesco in cases where a CAN/SDI service has been established experimentally, so that other SDI centres can be helped in the difficult task of evaluation. It should be noted, however, that this methodology is only given by way of an example and that other methodologies are cited in the bibliography.

The purpose of the proposed evaluation survey is to collect data on the characteristics of the user population and the service, and on the degree to which those users are satisfied with the service provided. It also sets out to make a precise quantitative evaluation of the service rendered by SDI, by comparing the cost of the service with the market value of the time-saving that users will have achieved in their documentary search activities by calling on a SDI service.

There are two important aspects to the proposed evaluation method: (1) the evaluation is carried out by questionnaire; and (2) the results are analysed preferably by a qualified evaluator working independently of the documentation centre responsible for the SDI service. These two aspects are examined in turn below.

4.1 Questionnaires

By examination of the replies obtained from a set of three questionnaires, an attempt is made to obtain a reasonable quantitative view of the value of the service rendered by the SDI. An English version of these questionnaires was used in Argentina and India, and the results of the evaluation and an improved version of the questionnaires taking account of the experience gained are given in the annexes. Figure 3 shows an outline schedule for the dispatch, collection and analysis of the questionnaires, based on the evaluation of an experimental phase involving one year's service. The details can obviously be adapted to other situations.

Questionnaire No. 1, which is concerned with the user's situation and documentary practice prior to SDI, is sent out just before the service commences.

Questionnaire No. 2 sets out to obtain data on the characteristics and impact of the service at the half-way point in the experimental phase (which is assumed in this case to be equal to five and one half months). It is sent out one month before this first half of the phase and is collected one month after. Users accordingly have two months in which to think about it, and this enables them to compile short series of statistics in reply to some of the questions, such as the average number of references per profile, the average quality of references and so on.

Questionnaire No. 3, which is similar to No. 2, is sent out towards the end of the experimental year, with provision again being made for the two-month period of reflection. It is assumed here that questionnaire No. 3 should be sent out after ten months of the service and collected after 12 months.

Each questionnaire is accompanied by a short letter designed to explain its purpose and to eliminate ambiguities. It also stresses the importance of complying with the deadline for returning the questionnaire.

The three questionnaires should be supplemented by the other items of information:

- the average hourly cost of the researcher's time, as derived from a short additional questionnaire;
- the breakdown of the computer costs and additional personnel costs of the SDI service during its first year of operation.

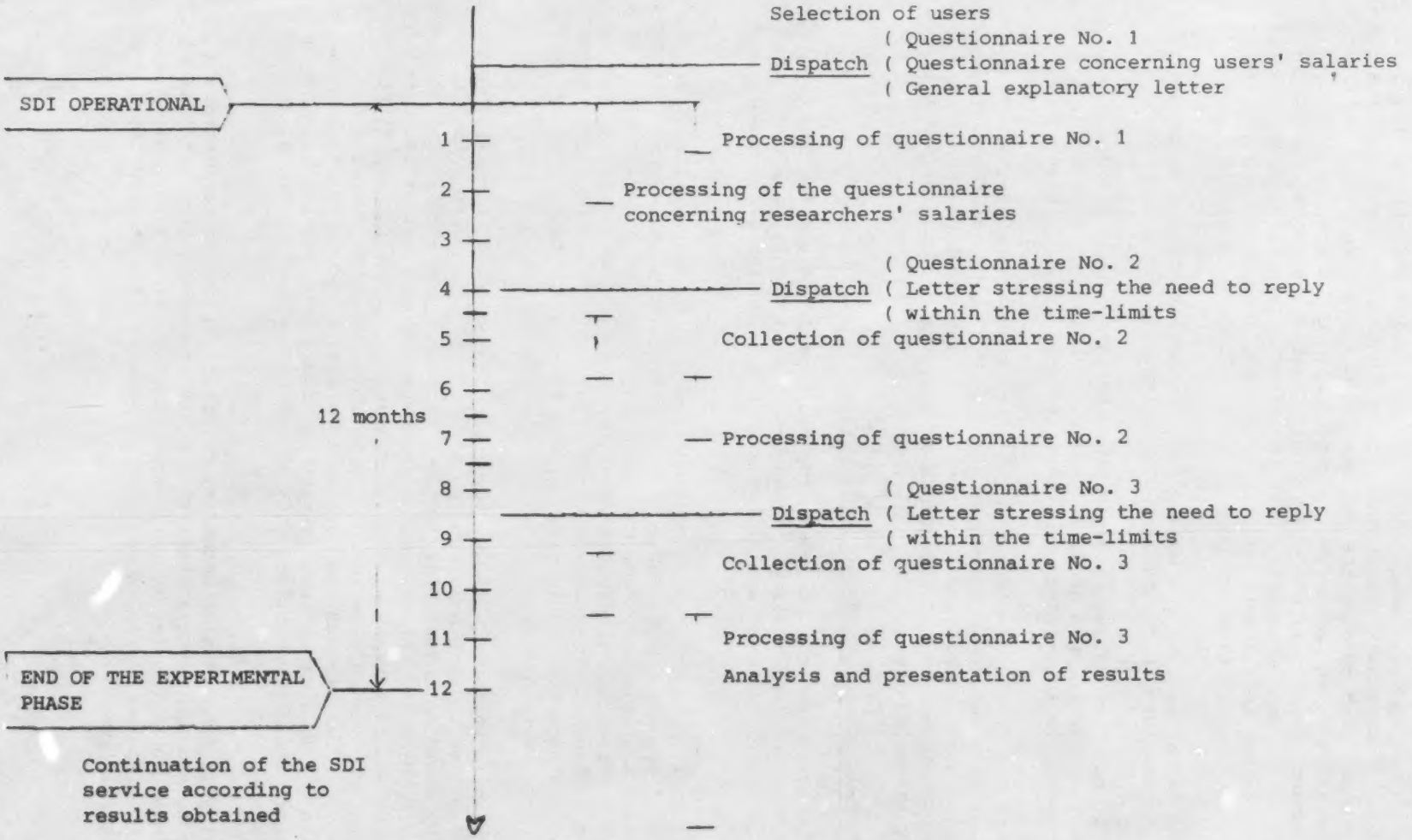
An example of the questionnaire concerning the user's salary is given below. This questionnaire is sent to all users at the same time as questionnaire No. 1. All due precautions to ensure anonymity should be taken in respect of this questionnaire. The user should be notified, in the accompanying letter, that in any event only the mean value of the salary for each type of employer organization will be used. In the specific case of students, the letter proposes that the remuneration quoted should be that normally offered in the first year of gainful employment in the student's chosen field.

It is to be noted that questionnaire No. 1 and that relating to salaries are to be completed by each user, whereas parts of questionnaires 2 and 3 relate to each profile, a distinction that does not need to be made if the rule "one user, one profile" is respected. Otherwise, it is necessary to add further explanations in the letters accompanying questionnaires 2 and 3 to make it clear that certain questions relate to a single profile and others to the SDI service in general. In any case, it is recommended that the SDI centre assign a single code number to each profile, and that this number should be entered on the questionnaires in the space provided to make it easier to monitor the replies. This approach also makes it possible to preserve the anonymity of the users' replies; in this case they will be duly notified in the accompanying letters, and the signing of the questionnaires will be made optional.

Two main difficulties are to be foreseen in the data collection: the long delays in obtaining replies to the questionnaires and the decreasing reply rate from one questionnaire to the following one. These two difficulties are interrelated and may be overcome by stressing the importance and anonymity of the evaluation, or by preselecting users likely to be motivated. This last approach should, however, be employed with care, since it has already been seen that it can jeopardize the efforts made to attract a large number of users and that it may limit the usefulness of evaluation as a forecasting tool.

Figure 3

OUTLINE OF THE STAGES IN THE EVALUATION OF A SDI SERVICE
(according to the methodology used for Unesco's pilot projects)



An insufficiently impartial evaluation or one based on a large number of "accommodating answers" would not be of much value and this is why it is useful to appoint an "independent" evaluator. In this case the questionnaires completed by the users are not sent through the SDI service but are forwarded directly and anonymously to the evaluator (whose address is provided in the corresponding accompanying letter). The documentation centre has everything to gain by agreeing not to see the individual replies to the questionnaires. The increased objectivity of the evaluation will be a cogent argument in following up the recommendations of the study. The centre will receive the final evaluation report which will contain all it needs to know.

One of two courses can be adopted to enable reminders to be sent out to any users who are late in returning the questionnaires:

- in the "triangular circuit", the evaluator sends the documentation centre a list of the code numbers of users who have replied. The centre can then keep its lists up to date and send out reminders to any users who have failed to reply;

- in the "L-shaped circuit", the users themselves undertake to inform the centre that they have replied to questionnaire No. 1, 2 or 3, as the case may be, using, for example, a short stencilled form prepared by the documentation centre. This approach relieves the evaluator of the burden of secretarial work for which the SDI centre is usually better equipped.

4.2 Processing and presentation of results

The evaluator will be chosen for his experience and for his qualities of objectivity and impartiality. A measure of familiarity with the users' environment would be useful since it would ensure that the replies and comments made by the users are more clearly understood.

It is from reading the questionnaires that the key will be found to processing the users' replies and comments. The resulting statistics will relate to: (a) the population of users and their behaviour in respect of document searches prior to SDI; (b) the profiles and the information supplied to users; and (c) the impact of SDI and the problems it presents to users. The results of the evaluation of the pilot projects in Argentina and India, set out in an annex to this study, may serve as an example of presentation.

It is recommended that not too much time be spent on analysing the distribution of the replies to each question, as the outcome of the evaluation has to be assimilated by decision-makers with little free time. For each reply, the main finding usually consists of the mean value, accompanied by a set of comments dictated by common sense. The basis of the calculation, i.e. the number of users or profiles entering into the calculation of the average, will be given for all important questions. The intermediate calculations will be performed to more decimal places than are needed in the final statistic. Since expressing a value as a round number is an irreversible operation, this should be done only once, at the end of the calculation.

It will sometimes be necessary to construct tables or histograms to make it easier to compare several statistics or to show the distribution of replies to a particular question (for instance the prices that different percentages of users would be prepared to pay to continue to benefit from the SDI service after the experimental phase).

EVALUATION OF SDI

ANONYMOUS SALARY SURVEY

1. Are you a student?

Yes ____

No ____

If so, refer to instructions attached.

2. Type of organization you are working in:

- government ..... ____)
- university ..... ____)
- private sector ..... ____)
- other (please do not specify) .. ____)

PLEASE DO NOT GIVE ANY OTHER DETAILS

3. Average monthly salary (before tax)

_____

_____

numerical value currency

A comparison of the replies made to certain deliberately identical questions, asked in questionnaires No. 2 and 3, will indicate trends in time of various specific aspects. In this connection, it has been found that a clearer picture is obtained of the problems involved (such as obtaining articles or having articles translated from foreign languages) from the replies to questionnaire No. 3. However, a relatively small response to this questionnaire, which might lead to the results being weighted in favour of the most highly motivated users, is to be feared.

In this type of survey it is quite common to find asymmetric distributions of replies, and it is possible that very few questions will result in a distribution of replies that seems to conform to a normal Gaussian curve. This problem, whether inherent in the question or linked to a too limited or non-representative sample, often makes it risky to use conventional tests to make a quantitative comparison of separate results (for instance replies to the same question at intervals in time, or results of different SDI services). If the presentation of the results alone does not suffice, the bibliography attached to this text gives the titles of a few works in which the relevant statistical problems are analysed in depth.

The length of time anticipated here for the processing of the questionnaires (one month each) is only indicative and depends on what information the evaluator is trying to bring to light. In particular, the time required to analyse the results obtained with questionnaire No. 2 and questionnaire No. 3 should not be underestimated. The schedule naturally presupposes that the letters accompanying the questionnaires have made it clear to the users that they must comply strictly with the deadlines for

returning the completed questionnaires. If the service complies fully with the "one user, one profile" principle, the time needed to perform ancillary calculations will be saved in the evaluation. If a large number of written comments are made by the users in questionnaire No. 3, a considerable amount of extra time will be needed to examine them.

The evaluator would be well-advised to start analysing the replies and drafting his calculations, observations and tentative comments, at least in manuscript, as soon as he has received a large enough number of replies to a particular questionnaire (No. 1, 2 or 3) since this will make it all the easier for him to draft his final report.

Lastly, the present methodology seeks to draw conclusion as to the cost-benefit ratio of SDI which is reflected, in particular, in a saving of time for the user (time which was formerly spent on manual document search and which can now be devoted to productive activities). This time-saving has to be expressed as a commercial value to be compared with the cost of the service, and for this purpose a survey needs to be made of the average salary cost (subsequently to be expressed on an hourly basis) of the researchers concerned (see subsection 4.1). In addition, the evaluator should, on completion of the experiment, end up with a figure for the various costs involved in the service itself (e.g. of subscription to the data base, the cost of using the computer and personnel costs) during the experimental phase of SDI. This will make it possible to calculate the average profile cost and to compare it with the annual time-saving effected by SDI, expressed as a commercial value. The relevant statistics can be expressed either as cost per researcher or as cost per profile.

This comparison will show the extent to which costs and benefits balance. The favourable effects of scale should always be borne in mind: all other things being equal, individual profile costs fall whenever the total number of profiles increases.

If, on completion of the experiment, the data on computer and personnel costs are not available, the evaluator will confine himself to showing the amount of time saved each year through SDI. The documentation centre will subsequently have to complete the calculations as soon as it obtains the relevant data.

If the SDI centre develops a mathematical model of costs (which can be gradually perfected and will be observation based), a more thorough-going analysis can be carried out. This model would use the available data on the service, such as the number of terms or references per profile, computer-time required, etc., and an attempt would be made to establish equations making it possible to predict the cost of a profile and, consequently, the profitability threshold of the service.

It should be recognized in conclusion that the figures expressing the costs and benefits covered by the present methodology, even if they are refined as far as possible, are by their nature approximate. It should also be stressed that the amount of time saved by the user is but one of the advantages of the service. The user may also be <u>better informed</u> through SDI, which may lead to a qualitative improvement in his work (for instance through his attention being drawn to a new and fruitful research topic). This aspect is covered by several questions relating to the users' degree of satisfaction, but it remains difficult to express it in quantitative terms. The results obtained during the evaluation of SDI should be carefully analysed and set in a realistic context so that they can serve as a means of improving the information services.

5. THE PLACE OF SDI IN CURRENT OR FUTURE COMPUTERIZED SERVICES

SDI is clearly not an end in itself, and the decision to set up such a service must be taken in a wider context which extends beyond the concerns of this guide. However, the question can rightly be asked whether under present circumstances and in the near future SDI at the local level will still be a valid information service in view of the recent expansion of international systems (especially systems of online interrogation) which offer universal bibliographic data bases for retrospective search and even for SDI.

Before this question is tackled a few essential principles need to be laid down:

- the choice of a particular information system should be made after mature reflection that takes into account the information policy of the country or the responsible body, which will have been framed in the light of the information needs and priorities;
- as far as possible, available resources should be used and care should be taken to avoid duplication or previous efforts;
- it is often recommended that national bibliographic data should first be effectively processed before planning the establishment of computerized services relying on universal bibliographic data bases.

In the light of these considerations, an evaluation of the main advantages of a local SDI service, in contrast to other information services using the universal bibliographic data bases, will provide a few pointers to decision-making. The local establishment of SDI does not necessarily preclude the use of outside services, meeting a variety of needs, often complement one another.

5.1 Establishment of SDI on a local basis

The local establishment of computer operations to conduct SDI on data bases, as envisaged in this guide, presents at least three main advantages:

(a) the country or organization concerned retains a measure of independence. Even though the data-processing equipment and data bases come from another country, the country concerned enjoys complete freedom in using SDI to process the themes of its choice. This specialization may also lead to a lowering of cost as compared with recourse to an outside bibliographical service, especially if a large number of SDI users can be obtained.

(b) the acquisition of new skills and technical knowledge in the field of computerized information systems represents an advantage for a country which does not yet have any experience in this area. The enhancement of available manpower skills in respect of both computer scientists and those responsible for profile design may also be a significant advantage.

Once a detailed grasp of software for SDI alone has been acquired, it may possibly be adapted, or the experience gained may be used for other purposes that are regarded as being desireable, such as the management of bibliographic data bases and even application outside of scientific or technical documentation such as administration, national statistics services of all kinds, and so on.

(c) Local establishment of computer applications may improve the return shown on equipment already purchased or leased, since the average hourly cost of computer time may decrease considerably as the daily number of house of use increases. (11)

Maximum use should consequently be made of a country's computers and the aim should be continuous operation wherever possible.

It has already been mentioned that as soon as the magnetic tapes have been processed in SDI every month, the computer is, in theory, free for all other uses.

Lastly, it should be stressed that the combination of the local processing of universal data bases in SDI with the interrogation of those same bases by a completely different channel for retrospective search is by no means rare. It has already been seen that the local exploitation of universal bases in SDI may prove profitable even though the local management of those same bases is often not feasible, and this fact may lead naturally to use being made of external services (e.g. the major hosts) for retrospective search. Furthermore, even if SDI and retrospective searches are carried out by local handling of universal bases, these two functions normally presuppose two independent software packages: SDI is ensured through the batch processing of new tapes, and the information on the tapes is subsequently incorporated in computerized retrospective files.

5.2 Use of international services

If a comprehensive view is to be obtained of the possibilities afforded by computerized documentation, account should be taken of the fact that universal data bases are becoming increasingly available throughout the world, particularly but not solely through systems of on-line interrogation.

The coverage of the bibliographical services offered by the major hosts to which access can be gained through teleprocessing networks is very large, as regards both available subject fields and the time period covered.

We shall not describe the commands or macroinstructions used for on-line interrogation of the major hosts. For this there are three reasons:

- these commands are in fact extremely straightforward, and, out of the 35 or so important instructions existing in most systems, only a fraction (around 10 to 15) are really in everyday use;

- the effect of these commands is very clearly explained in self-instruction manuals prepared by the relevant services of the major systems in question here;

- training sessions (for beginners and for personnel who have already received training) are held periodically and enable the on-line use of the systems to begin smoothly.

The failure rate of the major international systems has become extremely low considering the complexity of the link between the users and the system. To give an idea, so far as bibliographic systems are concerned, a serious breakdown may last half a day, although such breakdowns have now become altogether exceptional. A normal breakdown, which occurs very infrequently, lasts about half an hour on the average.

At the present time, for many countries the main obstacle impeding access to the major host computers is the inadequacy of international telecommunication facilities, or the high cost of using them for the transfer of information. Such obstacles are bound to become less insurmountable as new technologies are introduced and as new telecommunication infrastructures, such as co-operative networks specializing in the transmission of data and satellites, are gradually established.

In the meantime, steps should be taken to obtain on-line access to the major hosts only after feasibility studies have been carried out, and care should be taken to minimize the cost. As stated earlier, there always remains the possibility of having access to data bases for retrospective search or SDI by corresponding through the normal postal channels with the specialized centres.

In conclusion, two points need to be made concerning the effective use of international systems:

(a) If we consider the matter closely, we shall see that it is through careless usage that the major international documentation systems are described as operating "on-line". Only the actual framing of the question takes place "on-line". For reasons of cost, the references obtained are not requested via the data transmission network but by airmail from the host centre to the users. Long-range data transmission is, in fact, still a costly exercise. It is a major financial mistake to have a local terminal print out the 70-odd references compiled for a non-urgent problem, and the user will soon come to realize this from the size of his monthly bills. Further confirmation of the truth of this remark derives from the facts that the texts of abstracts are more and more frequently available in data bases and that both the time spent on line and the total number of characters transmitted are taken into account in charging for the terminal-computer link-up.

A request for interrogation in batch mode, which usually involves voluminous and non-urgent replies, is particularly suitable for SDI. In this case it is frequently possible to defer not only the reply but also the data processing which is performed in batches by the host computer during off-peak hours.

It would therefore be detrimental to propagate the image of instantaneous-response systems. Instead, it would be more effective to speak in terms of systems with good total response times. As a corollary to this, users should maintain a traditional far-sighted approach to satisfying their documentation needs, especially since a further period will elapse between the time when they are supplied with the reference and the time when they obtain the original document, and this will sometimes be aggravated by the need to have the original document translated.

(b) In view of the costs involved, it is imperative that on-line interrogation be conducted with the maximum efficiency and even speed. If good interrogation performances are to be achieved, it is necessary that, prior to the link-up with the host computer, the documentalist/intermediary give considered thought to the question to be asked, and furthermore that there be a minimum number of hours' interrogation every month, lest the exact details of certain instructions are forgotten. This pitfall can be avoided by requesting the intermediary to prepare for his personal use a short aide-memoire. Users are actually provided with this facility in some systems. It is also possible to step up the pace of work on the keyboard by simulating interrogations without actually going on line, until users reach the stage where they no longer have to "grope" for the characters on the keyboard. Lastly, questions should, as a rule, not be put in overly specific terms; otherwise there is a risk that no references will be retrieved.

5.3 Conclusion

It is not the purpose of this guide to recommend SDI as a universal solution. Each organization concerned must weigh up the advantages and disadvantages of the possible approaches to the current awareness service: computerized handling of data bases, at the national or regional level, local management of profiles to be processed outside, establishment of "manual" services based on traditional forms of secondary information, or a combination of these possibilities. What is important is that the different types of user have access to the current documentation needed for them to play their role in the development process.

ANNEX 1

MAIN RESULTS ACHIEVED USING THE CAN/SDI SOFTWARE PACKAGE IN ARGENTINA AND INDIA

1. INTRODUCTION

Two highly satisfactory experiments involving installation of the CAN/SDI software have taken place in Argentina (with several users in Bolivia) and India (with several users in Nepal and Sri Lanka), on the initiative of Unesco and the relevant national authorities, with training in the use of CAN/SDI being provided by CISTI (Canada Institute for Scientific and Technical Information).

Chemical Abstracts Condensates was the bibliographic data base chosen by both Argentina and India. The results to be presented here have emerged from analysis of replies to the questionnaires prepared to evaluate these projects which are themselves reproduced (in slightly improved form) in Annex 2. The evaluations were also the subject of detailed reports to Unesco. In the following pages an effort will be made to pick out the most significant results by highlighting:

1. the general characteristics and information-seeking procedures of users prior to the establishment of SDI;

2. the positive impact and technical characteristics of SDI;

3. the problems arising from SDI; and

4. lastly, a concise estimate of the cost and benefit of SDI for the data base under consideration here.

A few preliminary remarks need to be made in order to place in proper perspective the statistics given below.

The pilot experiments did not quite follow the general lines described in this guide. Conversely, actual examination of the results led to certain improvements being incorporated in the guide, in particular that it would be useful to carry out a preliminary study of the potential demand for information, making it easier to set up the service.

It should also be stressed that the researchers who took part in these experiments were not a statistically representative sample of all the researchers in the countries referred to. An attempt will nevertheless be made, solely for the sake of illustration, to compare the two experiments, the results of which are surprisingly similar.

Lastly, for the benefit of the specialists, a word can be said about the computer equipment employed:

- In Argentina: two computers were used, according to availability, an IBM 370/158 and an IBM 370/145. These computeres were located in Buenos Aires, as is the SDI centre (CAICYT), and belong to the Computation Centre of the Ministerio de Bienestar Social.

  In the final phase of the project, it was planned to utilize a new computer about which information was not yet available for this study.

- In India: an IBM 370/155 computer located at the Indian Institute of Technology (Madras) was used, the SDI centre (INSDOC) being in New Delhi.

2. GENERAL CHARACTERISTICS AND INFORMATION-SEEKING PROCEDURES OF USERS PRIOR TO THE ESTABLISHMENT OF SDI

2.1 GENERAL CHARACTERISTICS OF USERS

- In the two experiments the users are all researchers and are roughly of the same average age (44.4 years in Argentina; 42.9 years in India); the age distribution curves seem to be normal or Gaussian.

- University researchers (4 out of 10) outnumber researchers in private industry (3 out of 10) in Argentina; in the case of India, researchers belonging to government agencies (7 out of 10) are largely in the majority.

- The researchers are in charge of other researchers (5 out of 10 in Argentina and 7 out of 10 in India).

- The proportion of time devoted to active research is very much the same (57.0 per cent in Argentina; 59.7 per cent in India).

- Applied research predominates among users in both countries (4 to 5 cases out of 10) followed by basic research (2.5 to 3 cases out of 10).

- Fields of research can be placed in:

  . 12 major categories, the first three of which account for 50 per cent of the researchers, in Argentina;

  . 18 major categories, the first four of which account for 50 per cent of the researchers, in India.

2.2 INFORMATION-SEEKING PROCEDURES OF USERS PRIOR TO THE ESTABLISHMENT OF SDI

2.2.1 General observations:

- The quality of the information available was considered to be reasonably good in 7 to 8 cases out of 10 in both countries.

- In the case of a new problem, document search is of the exhaustive type in more than half the cases in both countries.

- Researchers have to go to a library to find interesting periodicals (5 times out of 10 in Argentina, 8 to 9 times out of 10 in India).

- For both groups of users, the increases desired in the documentation budget are very similar: + 68 per cent in Argentina, + 63 per cent in India, or rouhgly + 65 per cent in both countries.

- The researchers had very small prior knowledge of SDI before the beginning of the experiment: one researcher out of 4 (Argentina) or out of approximately 5 (India). Written information concerning SDI seems to be most effective (6 cases out of 10 for those having prior knowledge).

2.2.2 If the time allocated to documentation by researchers is broken down into four components as follows:

WA: time spent in getting to documentation centres ("walking time");

SC: time spent in leafing through journals to locate interesting articles ("scanning time");

OR: time spent in obtaining interesting articles that have been located ("ordering and/or photocopying time");

TR: time spent in absorbing documents ("true intellectual work of reading and assimilating");

the values set out in Table 1 below are seen:

Table 1

Total and relative components of documentation time prior to SDI

Country	Argentina, Bolivia	India, Nepal, Sri Lanka
Total documentation time (hours/week/researcher)	15.4	17.0
incl.: WA in %	9.1%	10.5%
SC in %	24.2%	24.7%
OR in %	10.2%	8.4%
TR in %	56.5%	56.4%
Number of replies	58 researchers	93 researchers

The following weighted values can now be calculated for all the 151 researchers:

Table 2

Average time devoted to documentation (in hours per week per researcher)

Argentina, Bolivia, India, Nepal, Sri Lanka; 151 researchers					
	$\overline{\overline{WA}}$	$\overline{\overline{SC}}$	$\overline{\overline{OR}}$	$\overline{\overline{TR}}$	Total
hours / week, researcher	1.6	4.0	1.5	9.2	16.3
% of total	10.0%	24.5%	9.0%	56.5%	100%

Extracting the ratios shown below, weighted for both countries, we arrive at the following:

$$100 \times \frac{\overline{\overline{SC}}}{\overline{\overline{SC}} + \overline{\overline{TR}}} = 30.2\ \%$$

$$100 \times \frac{\overline{\overline{TR}}}{\overline{\overline{SC}} + \overline{\overline{TR}}} = 69.8\%$$

The intellectual work of scanning and assimilating documentation thus breaks down as follows:

3/10 is spent on scanning for interesting articles,

7/10 is spent on assimilating the articles that have been chosen.

The assimilation of articles takes 2.3 times more time than it takes to scan for them.

In short, prior to SDI, documentation activities were broken down as follows:

- unavoidable material tasks:	- travelling:	10%	19%
	- ordering articles:	9%	
- intellectual tasks:	- scanning:	24.5%	81%
	- reading and assimilating:	56.5%	

It will be noticed that apparently little use is made of photocopying either in Argentina or in India. "OR" time, corresponding to "ordering and photocopying", is therefore used mainly for ordering articles.

The approximately 16.3 hours per week spent, on the average, on documentation are, for the most part (81 per cent), devoted to the "noble" aspect of that activity, i.e. finding interesting articles and, especially, assimilating them.

2.2.3 The number of primary and secondary journals read by researchers is shown in Table 3.

Table 3

Weighted averages of numbers of journals per researcher

Argentina, Bolivia, India, Nepal, Sri Lanka	
Primary journals/researcher:	16.0 (sample: 162 replies)
Secondary journals/researcher:	2.9 (sample: 146 replies)
Total:	18.9

Consultation of these 18.9 journals requires, as we have ssen, about 16.3 hours per week per researcher. The corresponding "average speed" is:

1.2 journals per hour

all phases (WA, SC, OR, TR) included, and all types of journals (primary and secondary) considered together.

As far as secondary journals are concerned, 2/3 of researchers in Argentina and 3/4 in India scan the hard-copy version of Chemical Abstracts "manually".

3. POSITIVE IMPACT OF SDI IN ARGENTINA AND INDIA

The information collected will be set out here under the following three headings:

- establishment and characteristics of the average profile;
- impact of SDI in satisfying documentation needs;
- time-saving resulting from SDI.

3.1 ESTABLISHMENT AND CHARACTERISTICS OF THE AVERAGE PROFILE

3.1.1 ESTABLISHING THE PROFILE

- Establishing the profile and making any necessary corrections require practically the same expenditure of time in the two projects on the part of the researcher: 3.30 hours per profile (Argentina) and 3.38 (India).
- The initial profile is satisfactory in about one instance out of two (55.6 per cent of cases in Argentina; 46.6 per cent in India).
- For the 50 per cent or so of profiles requiring subsequent changes, an average of 1.2 correction sessions per profile is needed in Argentina and 1.6 in India.
- These corrections concern descriptors in about half of the cases.

3.1.2 CHARACTERISTICS OF THE AVERAGE PROFILE

Profile characteristics can be reduced to three values:

- the number of search terms per profile;
- the number of references per profile per magnetic tape processed;
- the content of the retrieved references.

3.1.2/1 NUMBER OF SEARCH TERMS PER PROFILE

It can be observed that after a time the profiles tend to be more cautiously designed.

On the other hand, this is one of the rare points where there is a sharp distinction between Argentina and India, as indicated by Table 4 below.

Table 4

Average number of terms per profile

	After 6 months of SDI	After 11 months of SDI
Argentina	12.5	11.8
India	32.1	25.5

Table 4 shows that the profiles in India contain (after six months) 2.6 and (after 11 months) 2.2 times more search terms on average than those in Argentina. It is true, though, that the search themes are also different for the two countries. Moreover, it is to be observed that for both countries, the values are distributed asymmetrically. This corresponds, in fact, to a general observation which is valid for all SDI profiles: no Gaussian distribution presents itself here.

3.1.2/2 NUMBER OF REFERENCES PER PROFILE PER MAGNETIC TAPE PROCESSED

The findings are outlined in Table 5.

Table 5

References/profile/processed tape (mean)

	After 6 months of SDI	After 11 months of SDI
Argentina	21.1	18.5
India	51.7	49.3

At first glance, it might seem that the 100 per cent difference observed as regards the references retrieved per profile is attributable to the identical difference observed for the number of terms/profile (Table 4). The experience of the French Petroleum Institute seems to indicate that this is not the essential factor. It is, in fact, possible to increase the number of terms in a profile not in order to expand a search field and thus obtain a greater number of references, but rather to achieve greater accuracy and thus reduce the number of references.

In our opinion, the real reason for the disparity is to be found in the structure of the data base. The Chemical Abstract Condensates are, as we know, divided into odd- and even-numbered issues. Normally, a search theme is slanted towards one of the two series of issues, seldom to both simultaneously.

Be that as it may, it should be noted that, if we have correctly understood the matter, the method of operation was not identical in the two countries:

- In Argentina each tape was processed individually upon reception and thus users were provided alternately with the data obtained from a tape in an odd-numbered issue and those obtained from an even-numbered issue.

- In India, on the other hand, pairs of tapes, i.e. one tape from the odd-numbered series and one tape from the even-numbered series, were processed simultaneously.

The result is that in Argentina, for every other tape, the number of references has to be low or almost zero. In India, however, the printout as received by the researcher every two weeks always contains references inasmuch as two tapes belonging to odd- and even-numbered series are processed simultaneously. These considerations are enough to explain why the figures for one country are twice as high as in the other,

without the number of terms per profile being necessarily involved. The ratio is, nevertheless, slightly more than 2, as indicated below:

	After 6 months of SDI	After 11 months of SDI
India / Argentina	2.4	2.7

This might suggest that the average profile in India supplies a somewhat greater number of references than the average profile in Argentina. Here, too, the nature of the search themes comes into play.

3.1.2/3 CONTENT OF RETRIEVED REFERENCES

The content of references is initially broken down into five categories.

In the "rapid method", the percentages are calculated by averaging the values ascribed to the different profiles independently of the number of references/profile of each profile considered.

In the "weighted method", on the contrary, the number of references/profile for each profile are taken into consideration. Although this method involves more work, the results thus obtained give a more accurate picture of the average quality of the references.

We shall take a quick look at the results provided by these two methods.

3.1.2/3.1 The results obtained by the "rapid method" are shown in Table 6.

The number of replies after 11 months of SDI in Argentina is relatively small. By contrast, in the case of India, the number of replies is very much the same after six months of SDI and 11 months of SDI. A decrease of some 7.6 percentage points in the frequency of irrelevant references (from around 34.5 per cent to around 26.9 per cent) is to be observed.

The references for immediate use - again in the case of India - increase by about 7.5 percentage points, progressing from about 20.3 per cent to 27.8 per cent.

It will, of course, be noted that the ambiguous references (category 4), those whose "interest cannot be determined because the reference does not provide enough details", are not very numerous (around 10 per cent) and correspond to a provisional state.

The law of numbers suggests that category 4 will disappear and be absorbed by the four remaining categories in proportion to the values initially observed, after all the references have been checked in the Chemical Abstracts ("on paper"). The distribution of category 4 among the four remaining categories makes it possible to show the overall documentation relevance as defined by the sum of the new values obtained for categories 1, 2 and 3.

Table 6

Types of references ("rapid method")

Types of references	ARGENTINA		INDIA	
	After 6 months of SDI	After 11 months of SDI	After 6 months of SDI	After 11 months of SDI
	Frequency in %	Frequency in %	Frequency in %	Frequency in %
1. For immediate use	24.3	28.8	20.3	27.8
2. Of interest but already found elsewhere	14.5	24.1	9.8	10.5
3. Of interest but of no immediate use	20.7	17.3	24.6	24.1
4. The interest cannot be determined because the reference does not provide enough details	13.1	7.3	10.8	10.7
5. Irrelevant	27.4	22.5	34.5	26.9
Number of replies	70	21	70	73

This gives the following results:

Table 7

Overall documentation relevance

	ARGENTINA		INDIA	
	6 months	11 months	6 months	11 months
References relevant in the broad sense	68.4%	75.8%	61.4%	69.8%
Non-relevant references	31.6%	24.2%	38.6%	30.2%

3.1.2/3.2 The "weighted method" takes into consideration the number of references for each profile. After distributing category 4 among the four remaining categories, we arrive at the figures set out in Table 8 based on the final stage (11 months of SDI). It can be checked by comparison with Table 7 that the weighted method gives individual values that are considerably different from those produced by the rapid method (5 to 15) points in respect of the overall percentage of relevant references) but that general trends remain the same.

Table 8

Content of references (weighted method) after redistribution of category 4, after 11 months of SDI

	ARGENTINA		INDIA	
	Frequency in %	%	Frequency in %	%
1. For immediate use	27.1	)	25.9	)
2. Of interest but already found elsewhere	16.0	) 60.6	11.2	) 65.2
3. Of interest but of no immediate use	17.5	)	28.1	)
4. Irrelevant	39.4	39.4	34.8	34.8

Table 8 calls for two comments:

- The SDI centres conserve a certain safety margin in the form of non-relevant references. It was stressed elsewhere (see subsection 3.5) that this was the way it should be since it places the real interest of the researcher in the forefront.

- The percentage probability that a reference that is relevant in the broad sense may be of immediate use to the searcher is:

. for India: $100 \times \frac{25.9}{65.2} = 39.8\%$

. for Argentina: $100 \times \frac{27.1}{60.6} = 44.8\%$

It is to be noted that a reference that is relevent from the broad point of view of documentation work has about four chances in ten of being of immediate use to the researcher.

This comparison between the point of view of a documentation centre and that of researchers is so seldom made in the relevant literature that it seemed worth while to call attention to it.

In 75 per cent of the cases in Argentina and in 70 per cent of the cases in India, a minimum of 60 per cent of relevant references was readily obtainable. In one out of every four cases in India they were obtained with some difficulty and in one case out of about five in Argentina they were not obtainable at all.

The great majority of researchers (94 per cent in Argentina, 85 per cent in India) several times made a comparison between computer-produced references and secondary journals produced "on paper".

The profiles are regarded as reasonably exhaustive or, in other words, they provide in most cases all the important references in the searchers' fields, as is shcwn below:

Affirmative replies, in % of searchers	After 6 months of SDI	After 11 months of SDI
Argentina	78.3%	75.0%
India	68.2%	70.5%

Finally, the readability of the computer-produced references was considered good in 62.8 per cent of cases in Argentina and 61.1 per cent of the cases in India.

3.2 IMPACT OF SDI IN SATISFYING DOCUMENTATION NEEDS

The documentation impact of SDI involves five aspects.

3.2.1 The feeling of being better informed, thanks to SDI, is expressed by 8-9 researchers out of 10. The affirmative answers remain about the same after 6 months and 11 months of SDI for Argentina (86.9 per cent and 87.5 per cent respectively); they increase slightly in the case of India (74.3 per cent and 81 per cent respectively).

3.2.2 The fact that it is unnecessary, thanks to SDI, to duplicate scientific research already conducted elsewhere appears to be an aspect of lesser importance (38.3 per cent affirmative answers for Argentina, 16 per cent for India). This can be explained by the very nature of research in which each researcher seeks to make a contribution, necessarily based upon the findings of others. There is, then, an unavoidable overlapping between new research and that done earlier.

3.2.3 A considerable saving of time for research in progress, made possible by SDI, is recognized by researchers in 4-5 cases out of 10 (48.3 per cent affirmative responses for Argentina, 40.6 per cent for India).

3.2.4 The detailed impact of SDI on research in progress[1] was evaluated as follows (the percentages represent the affirmative answers as against the number of answers recorded after 6 months of SDI):

- direct, general assistance of SDI:

  Argentina: 87.3% India: 79.7%

- SDI provision of "state-of-the-art" reviews:

  Argentina: 67.6% India: 58.1%

- SDI provision of details on experimental and/or industrial methods:

  Argentina: 73.3% India: 35.1%

- SDI provision of patents showing that there are good grounds for the work in hand:

  Argentina: 29.6% India: 39.2%

- SDI provision of patents hampering current work considerably:

  Argentina: 9.9% India: 13.6%

3.2.5 The overall degree of reliance on SDI is seen from the following:

- SDI has become the sole documentation source:

  Argentina: 0% (not one affirmative reply)

  India: 3.80%

- SDI is an extremely useful supplement to "manual" documentation in general:

  Argentina: 50.0% India: 51.9%

- SDI is frequently useful:

  Argentina: 43.8% India: 22.8%

- SDI is occasionally useful:

  Argentina: 6.2% India: 17.1%

- SDI is of no use; traditional methods are preferable:

  Argentina: 0% India: 4.4%

(1) The relevant question has been slightly recast in the present version of the questionnaires.

3.3 SAVING OF TIME MADE POSSIBLE BY SDI

The value $\overline{GT}$ designates the average time-saving per working week per researcher. It represents the difference between the total time devoted to documentation before SDI (or TTDOC) and with SDI. The values are summarized in Table 9 below:

Table 9

Values in hours/working week per researcher

	Argentina	India	General weighted average
$\overline{GT}$ (weighted average) 6 months and 11 months of SDI	4.6	5.3	5.0
$\overline{TTDOC}$	15.4	17.0	16.3
ratio $\frac{\overline{GT}}{\overline{TTDOC}}$	0.302	0.309	0.307

In this table, the value 5.0 is based on 95 researchers (aggregate after 6 and 11 months of SDI), and the value 16.3 is based on 151 researchers (before SDI).

We can therefore take as a rule of thumb:

$$\overline{GT} \sim 0.3 \times \overline{TTDOC} \qquad \left[\frac{\text{hours}}{\text{week/researcher}}\right]$$

This rule could be valid for other services using the same data base (CAC) in so far as the various components of $\overline{TTDOC}$ (WA, SC, OR, TR, see Table 1) are, in relative value, analogous. Clearly this rule can be applied only to users whose search theme is reasonably well covered by the base.

In the case of Argentina, more than half of the users received information based on more than one profile (1.7 profiles per researcher). It is very interesting to note that, according to the relevant statistical tests, the number of profiles per researcher does not seem to exercise any significant influence on the amount of time saved per profile. This finding could not be established in the case of India where all the users were, in principle, entitled to only one profile.

4. PROBLEMS ARISING FROM SDI

These problems are of two kinds:

- availability of journals
- translation of articles.

Table 10 below outlines the findings.

Table 10

Influence of SDI on the problems of access to journals

Stage		Argentina	India
Before SDI	Researchers already experiencing some difficulties: (number of replies)	61.7% (45)	58.9% (73)
After 6 months of SDI	Researchers who consider that SDI gives rise to new problems: (number of replies)	20.3% (43)	38.8% (67)
After 11 months of SDI	Researchers who consider that SDI gives rise to new problems: (number of replies)	71.4% (14)	68.1% (72)

The example of India - for which the number of replies remains fairly constant - shows that after 11 months of SDI, two researchers out of three feel that SDI gives rise to new problems of access to journals as compared with one out of three after six months of SDI.

Table 11 below outlines the findings concerning the problem of having articles translated from foreign languages.

Table 11

Influence of SDI on problems relating to the translation of articles

Stage		Argentina	India
Before SDI	Researchers already experiencing some difficulties: (number of replies)	31.1% (45)	50.7% (75)
After 6 months of SDI	Researchers who consider that SDI gives rise to new problems: (number of replies)	6.8% (44)	29.2% (65)
After 11 months of SDI	Researchers who consider that SDI gives rise to new problems: (number of replies)	40.0% (15)	44.4% (63)

Again there appears to be greater awareness of the problems after 11 months of SDI than after six months.

This calls for two comments.

The appearance of problems of access to the journals reported by SDI and concerning the translation of articles reported by SDI, is an integral part of the SDI picture. It simply reflects the following two aspects of a situation arising from (or made possible by) SDI:

- the documentation of the researcher becomes more complete and he discovers journals of which he had been unaware or which had not been processed up to then and which prove useful for his work;

- the researcher's documentation becomes more international, which in turn, explains the problems concerning the translation of articles of interest, brought to light by SDI.

The second comment stems from the first. Computerizing the selection of documents cannot solve the other major problems of documentation (including the provision of journals and the possibilities of translation). The "documentation chain" constitutes a whole of which SDI represents only one link and it is thus essential to take account of all aspects - even the most conventional - of that chain.

5. COST AND PROFITABILITY OF SDI

5.1 INDIA

By way of introduction, let us recall that the hourly cost of a computer depends on the number of shifts assigned to it, or, in other words, on the computer's work-load. The bulk of the rental cost of the computer relates to the cost of the work done by the first eight-hour shift. It is practically negligible for the other two possible shifts. We should also mention that the owner of the computer is free to charge the average cost or the marginal cost for the second and third shifts, provided there are suitable reasons for doing so.

In India, the expenditure incurred by the organizers for the first year of SDI was as follows, expressed in rupees (Rs):

Computer:	Rs118,000
Personnel:	Rs 32,000
Total:	Rs150,000

The number of profiles varied between 100 at the start of the experiment and 150 at the conclusion, giving an average of 125 profiles (corresponding to the same number of users, each user being entitled to only one profile). The cost of the average profile is therefore Rs1,200.

It should be remembered that the profiles involved here are matched against magnetic tapes from the even- and the odd-numbered series (see page 35), and that the tapes are processed once every two weeks.

The average salary of a researcher in India can vary between Rs500 and Rs3,000 per month, i.e. between Rs6,000 and Rs36,000 per year.

We shall make our calculations on the basis of a lower limit of Rs10,000 (for an average researcher) and an upper limit of Rs20,000 (for an experienced researcher) per researcher per annum.

On the basis of 1,800 working hours per year (a figure applicable to several industrial countries),(1) we arrive at the following range of hourly costs:

(1) See for example, the weekly economic journal Valeurs Actuelles, Paris, 7 August 1978, p. 19.

5.555 Rs/hour (average researcher)

11.111 Rs/hour (experienced researcher).

As was seen earlier (Table 9), the overall time devoted to documentation was 17.0 hours per week. That figure in relation to a five-day working week of eight hours per day is equal to 42.4 per cent of the work-time of the researcher, i.e. 763 hours/year on the basis of 1,800 hours of work per annum.

SDI has made it possible, as we have seen (Table 9), to achieve a saving of 30.9 per cent of overall documentation time, i.e.:

annual time-saving = 9.309 x 763 hours/year

= 226 hours per researcher/year.

This saving in time can be converted into monetary units by using the range of hourly costs per researcher given above. Thus:

$$236 \frac{\text{researcher hours}}{\text{year}} \times \frac{\text{Rs } 5.555}{\text{hour}} = \frac{\text{Rs1,311}}{\text{year}}$$

$$\text{and } 236 \frac{\text{researcher hours}}{\text{year}} \times \frac{\text{Rs11.111}}{\text{hour}} = \frac{\text{Rs2,622}}{\text{year}}$$

These figures may be compared with the average cost of the service per researcher in India, which, as noted earlier, amounts to Rs1,200/year (cf. above).

We find that, even at the lower rung of the salary scale (Rs10,000/year), the operation is sound (it results, in fact, in a slight profit: 1,311 - 1,200 = Rs111/year savings). At the top rung of the salary scale, corresponding to an experienced researcher (Rs20,000/year), the operation is highly profitable:

2,622 - 1,200 = Rs1,422/year saving per researcher (per profile), i.e. the net gain made possible by SDI per year and per researcher, a net gain which stays in the country concerned and which supports the work of national specialists and assures the profitability of the country's computers.

Admittedly, the cost of the subscription to the data base does not appear here. In the case under consideration, the subscription to the CAC base amounted to Rs12,000 (about US $1,340 in 1976), which represented a preferential rate. This accounted for less than 8 per cent of total expenditure and did not perceptibly affect the results of the evaluation. Of course in each particular case this question needs to be reconsidered. It should be pointed out that, where certain data bases are concerned, the cost involved is necessarily the combined cost of journals on paper and magnetic tapes. In such cases there is no rational criterion for dividing up the overall cost.

Lastly, the favourable effects of scale should not be overlooked. Thus, in annual spending, the cost of the subscription to the tapes is generally independent of use to a large degree, and the cost of processing them may be relatively unaffected by the number of profiles involved. The relative cost of a profile tends to be in inverse proportion to the total number of profiles.

5.2 ARGENTINA

In the case of Argentina, we received no breakdown of expenditure. It will therefore be advisable to complete the calculations locally as soon as these data become available.

There is, notwithstanding, good cause to be just as optimistic and for two reasons:

- the average number of terms/profile is smaller in Argentina than in India and this number directly affects the amount of time spent on (and hence the cost of) the data processing needed for searching through the expression D x T (see subsection 3.4)

  where D is the average number of documents/magnetic tape and T is the total number of search terms, the product of the average number of terms/profile and the total number of profiles; moreover, D is high in the case of Chemical Abstracts Condensates (in the neighbourhood of 8,000 documents/tape);

- furthermore, the salary cost of a researcher is higher in Argentina(12).

It is to be recalled, however, that, in contrast with the procedure adopted in India, Argentian researchers could receive information corresponding to more than one profile and this tended to raise the cost of data-processing per user. This seems to have been offset, however, by the fact that the amount of time saved by each user in regard to documentation also increased in the same proportion.

5.3 AVERAGE VALUE OF THE SDI SERVICE

The estimates made by the users of the average value of the SDI service prove to be, relatively speaking, very similar for the two experiments under study here, as shown in the synoptic table below.

Table 12

Average value of SDI service, from the point of view of the users, expressed in US dollars per year per profile

	Argentina	India
After 6 months of SDI:	US $71.86	US $55.11
After 11 months of SDI:	US $74.21	US $59.20
Average (unweighted) by country:	US $73.04	US $57.16
Combined (unweighted) average:	US $65.10	

It by no means follows that one can ask for the equivalent of US $65 per year per profile in each of these countries. Any possible price policy must take into careful account a number of factors and, in particular, the type of user. In the final analysis, any decision about price policy will have to be made by the organization responsible and will depend on the documentation policy that is followed in the medium and long term.(1)

(1) The distributions of perceived value are highly asymmetrical, the value appearing most frequently (or "mode") being less than the average value. The histogram of the distribution should therefore be closely examined so as to draw useful conclusions regarding a subsequent price policy. It is to be noted that, of all the parameters covered by the evaluation, only the users' ages seem to conform to a Gaussian curve.

6. RETROSPECTIVE SEARCH

Regarding the period of time which should be covered by retrospective search, this should correspond, according to the users, to an average of five years (for Argentina) and seven years (for India) or a combined average of six years.

This information may be pertinent to use of the CDS/ISIS component of the CDS/ISIS - CAN/SDI software package offered by Unesco and, in particular, to development of a version of CDS/ISIS whose capacity could be enlarged through software modifications which make it possible to conduct retrospective searches in the major data bases.

Elsewhere we have stressed the fact that a subject-field should be of great interest in terms of retrospective search requests in order to warrant the local management of an entire data base in that field. We likewise considered it worth while to suggest that priority areas of documentation for data base management and retrospective retrieval are often, but not exclusively, of national origin. True, the choice is a vast one, but it nevertheless falls within the province of the relevant national authorities.

Such are the main results to emerge from our analysis of the questionnaires. Other information which was of less importance in the context of the present study has, for reasons of space, not been included.

ANNEX 2

QUESTIONNAIRES USED FOR THE EVALUATION OF EXPERIMENTS
WITH SDI IN ARGENTINA, INDIA AND NEIGHBOURING COUNTRIES

These Questionnaires have been prepared by the Institut Français du Pétrole, under UNESCO contract.

- SDI EVALUATION -

You will receive a total of 3 questionnaires :

No. 1 before any use of computerized documentation systems for Selective Dissemination of Information (SDI);
No. 2 5.5 months after implementation of SDI Services;
No. 3 11 months after implementation of SDI Services.

This questionnaire is : QUESTIONNAIRE No. 1

Each SDI user is requested to fill out one Questionnaire n° 1

Please do not use this column

1	Please write here your personal CODE (lowest value of your codes if you have more than one profile) ☐	
2	Age*: 25 or under ☐ 41-45 ☐ 26-30 ☐ 46-50 ☐ 31-35 ☐ 51-55 ☐ 36-40 ☐ 56-60 ☐ 61 or over ☐ * at last birthday	
3	Educational level (highest qualification obtained): - Diploma or vocational certificate ..... ☐ - Undergraduate degree ..... ☐ - Masters degree ..... ☐ - Doctorate ..... ☐ - Other (specify) ..... ☐	

Questionnaire No. 1 - page 2

4	Type of organization you are working in: government ☐ university ☐ private sector ☐ other (please specify):	
5	Present position: - individual researcher ..... ☐ - member of a research team ..... ☐ - in charge of researchers ..... ☐ - other (please specify):	
6	Percentage of time devoted to active research (as opposed to general management, administrative work, etc.) 0-20% ☐ 21-40% ☐ 41-60% ☐ 61-80% ☐ 81-100% ☐	
7	Field of activity (in very general terms only: preparation of polymers on a semi-industrial level; biochemistry applied to vitamins; etc.):	
8	Is your research considered as: - fundamental (adding to scientific knowledge in general) ..... ☐ - applied (intended to produce knowledge of pratical value) ..... ☐ - experimental development (solution to a specific problem, including prototypes and pilot plants) ..... ☐ - other (please specify: long-range planning, economic studies, etc.):	

Questionnaire No. 1 - page 3

9	How do you consider that you are presently informed on new work in your field of activity? - in an exhaustive manner ... ☐ - reasonably well ..... ☐ - poorly ..... ☐ additional comments as appropriate (e.g. poorly because of lack of time, etc.):	
10	When faced with a new problem, how do you do your literature search? - first try to have an exhaustive view ..... ☐ - look only at the most recent literature ..... ☐ - no literature survey at all because literature is non-existent ..... ☐ because it is too time-consuming ..... ☐ Other comments as appropriate:	
11	How is documentation organized in your establishment? - all (or practically all) the interesting periodicals arrive on your personal desk by circulation ..... ☐ - half (or about half) of them arrive on your desk, the other half must be consulted in some other place ..... ☐ - all (or practically all) of them must be consulted in another place ..... ☐ Additional comments as appropriate:	

Questionnaire No. 1 - page 4

12 Please give the titles (abbreviated titles will do) of the technical periodicals you scan very regularly (excluding Abstracts and Lists of titles, see next question).

1	9
2	10
3	11
4	12
5	13
6	14
7	15
8	16

(Give the total number if more than 16: with the details for the first 16). Note that the value 16 has no special meaning. ----------

13 Please give the titles of Abstracting journals and Lists of titles you scan very regularly: e.g.: Chem. Abstr.

1	5
2	6
3	7
4	8

(Give the total number if more than 8: with the details for the first 8). Note that the value 8 has no special meaning. ----------

14 How much time do you spend for your scientific documentation as broken down hereunder in hours/working week:

- walking to librairies ..... [ ]
- scanning for interesting articles ..... [ ]
- ordering, copying or photocopying ..... [ ]
- true intellectual work of reading and assimilating ..... [ ]

58 total in hours/week: [ ]

This question is important for future evaluation. Please take the time to give the best estimate (annual average) you can give here.

Questionnaire No. 1 - page 5

15	In your opinion, the documentation budget in your organization - is too low and should be increased by ..... % if possible - is presently adequate for your needs ☐ - is too high and might be reduced by ..... % Special comments as appropriate (not enough photocopying machines, etc.):	
16	Do you already know computerized documentation systems: Yes ☐ No ☐	
17	If your answer to question 16 was yes, how were you informed about computerized documentation systems? - by written publicity on them Yes ☐ No ☐ - by a demonstration Yes ☐ No ☐ - because you already use a computerized documentation system Yes ☐ No ☐ - other comment as appropriate:	
18	Date: Signature:	

Your cooperation is greatly appreciated. Please carefully check to make sure that all the questions have been answered: this is most important for proper evaluation and improvement of our documentation services. Be sure to keep trace of your CODE(S) for the other two questionnaires coming in 5.5 and 11 months. Thank you.

These Questionnaires have
been prepared by the
Institut Français du Pétrole,
under UNESCO contract.

- SDI EVALUATION -

QUESTIONNAIRE No. 2

This questionnaire is concerned with the first 5.5 months of use of computerized documentation systems for Selective Dissemination of Information. (SDI)

One copy of Questionnaire N° 2 is being provided for each SDI profile. If you receive service based on more than one profile please complete the first nine questions on a separate questionnaire for each profile. Questions 10-21 need only be answered once by each SDI user.

		Please do not use this column
1	Please write here your personal CODE or the code of the concerned profile: ☐	
2.1	- Your total number of profiles for SDI Service is :	
2.2	- Give here your other codes if you have more than 1 profile:	
3	How much time in hours did you devote to your profile for the initial profile + subsequent changes (if any)? 1 hour ☐ ; 2 hours ☐ ; 4 hours ☐ ; 8 hours ☐ ; more (specify) :	
4	How many times did you make a change in your profile?	
5	What are the changes you made in your profile? adding — or — suppressing ☐ descriptors ☐ ☐ authors'names ☐ ☐ section numbers ☐ ☐ other data elements ☐	

Questionnaire No. 2 - page 2

6 Have you the general feeling that your profile, after any possible changes you have made in it, is now reasonably exhaustive, i.e. all important references on your subject are retrieved?

Yes ☐ No ☐

7 What is the (approximate) number of terms in your profile ?

1-10 ☐ ; 11-20 ☐ ; 21-30 ☐ ;
31-40 ☐ ; 41-50 ☐ ; 51-60 ☐ ; more than 60: .....

8 What, on the average, is the total number of references retrieved for your profile per SDI issue (your best estimate) ?

1-10 ☐ ; 11-20 ☐ ; 21-30 ☐ ; 31-40 ☐ ;
41-60 ☐ ; 61-80 ☐ ; 81-100 ☐ ; 101-150 ☐ ;
151-200 ☐ ; more than 200 (specify): ...

9 How would you break down the retrieved references into the five following categories (in percentage)?

1 - of immediate interest: ..... %

2 - of interest but already found elsewhere: ..... %

3 - of interest but not for immediate use: ..... %

4 - the interest cannot be determined because the citation does not provide enough details:..... %

5 - of no interest, irrelevant: ..... %

★

10 Before the use of a computer in documentation, did you have any problems concerning journal articles that were not readily available to you:

Yes ☐

61

No ☐

★ NOTE : Starting from here, only one answer per SDI user is needed. Users having more than one profile need only to answer once (on the Questionnaire for the lowest profile code); they need not recopy their answers onto their subsequent copies of this Questionnaire.

11	Since you have been using computerized documentation, has your overall problem concerning availability of journal articles decreased ☐ ; stayed the same ☐ ; increased ☐ Please give details as appropriate:	
12	Before the use of computers in documentation, did you have any problems concerning translation of articles or books? Yes ☐ No ☐	
13	Since you have been using computerized documentation, has your overall problem concerning translation of articles decreased ☐ ; stayed the same ☐ ; increased ☐ Please give details as appropriate:	
14 14.1 14.2	Do you feel you spend less time for your scientific documentation ? Yes ☐ No ☐ If possible, give estimated times in hours/working week - before computerized documentation: _____ hours/week - with computerized documentation: _____ hours/week This question is important for the evaluation. Please take the time to provide the best estimates you can give here. (There is no obligation to use round figures if the quality of your estimates justifies the presence of decimals).	
15	Do you feel better informed with computerized documentation? Yes ☐ No ☐	

16 Has computerized documentation brought to your attention any papers that prevented duplicating research done elsewhere?

Yes ☐ No ☐

17 Has computerized documentation brought to your attention any papers that have saved a considerable amount of time in your own research?

Yes ☐ No ☐

18 Has computerized documentation brought to your attention any papers that have contributed directly to the problems you are dealing with?

18.1 Yes ☐ No ☐

In what manner?

- by providing a review of the state of the art in your field of work

18.2 Yes ☐ No ☐

- by giving you details on experimental or industrial methods (underline the appropriate adjective)

18.3 Yes ☐ No ☐

- by giving you important results and/or conclusions

18.4 Yes ☐ No ☐

- by revealing patents showing that your research direction is good

18.5 Yes ☐ No ☐

- by revealing patents showing that your field of research is already fully covered by patents

18.6 Yes ☐ No ☐

Questionnaire No. 2 - page 5

19 The readability of the references given by the computer is in your opinion:

good ☐

medium ☐

poor ☐

Additional comments as appropriate:

20 How much would you (or your organization) be willing to pay to have the computerized documentation service continued (values are per year and per profile)

numerical value	currency

Comments as appropriate:

21 Date: Signature:

Your cooperation is greatly appreciated. Please carefully check to make sure that all the questions have been answered: this is most important for proper evaluation and improvement of our documentation services. Be sure to keep trace of your CODE(S) for the last questionnaire coming after 11 months of experience with computerized documentation. Thank you.

These Questionnaires have been prepared by the Institut Français du Pétrole, under UNESCO contract.

- SDI EVALUATION -

QUESTIONNAIRE No. 3

- This questionnaire is for users having (for most of them) 11 months of experience with computerized documentation systems.
- There is an intentional repetition of some questions for the purpose of measuring the modifications since the previous questionnaire. Please attempt to answer these questions with full objectivity.

One copy of Questionnaire N° 3 is being provided for each SDI profile. If you receive service based on more than one profile, please complete the first eight questions on a separate questionnaire for each profile. Questions 9-21 need only be answered once by each SDI user.

		Please do not use this column
1.	Please write here your personal CODE or the code of the concerned profile: ☐	
2.	- Give here your other codes if you have more than 1 profile:	
3	When did your profile start? year: 19.. Jan ☐ Feb ☐ Mar ☐ Apr ☐ May ☐ Jun ☐ Jul ☐ Aug ☐ Sep ☐ Oct ☐ Nov ☐ Dec ☐	
4	Do you think that your profile gave you all the important references in your field of interest: Yes ☐ No ☐	
5	Did you compare computer output and paper issues when available? - never ☐ - once ☐ - a few times ☐ - for about (specify) ... % of the outputs.	

Questionnaire No. 3 - page 2

6 It is assumed that the general characteristics of your profile are unchanged since Questionnaire No. 2. If important changes had to be made since then, please indicate here the new average characteristics:

6.1 - Number of terms in your profile

1-10 ☐ ; 11-20 ☐ ; 21-30 ☐ ;
31-40 ☐ ; 41-50 ☐ ; 51-60 ☐ ; more: .....

6.2 - Total number of references per SDI issue (your best estimate for the profile in question) :

1-10 ☐ ; 11-20 ☐ ; 21-30 ☐ ; 31-40 ☐ ;
41-60 ☐ ; 61-80 ☐ ; 81-100 ☐ ; 101-150 ☐ ;
151-200 ☐ ; more than 200 (specify): ...

7 From your 11 months of experience with a computerized documentation system, how would you now break down the retrieved references into the following five classes for the profile in question (in percentage)?

1 - of immediate interest ..... %

2 - of interest but already found elsewhere ..... %

3 - of interest but not for immediate use ..... %

4 - the interest cannot be determined because the citation does not provide enough details .... %

5 - of no interest, irrelevant ..... %

8 Did you have any difficulties in elaborating your profile? More specifically:

8.1 - Extracting all the interesting references from the database was :

easy ☐ difficult ☐ impossible ☐

8.2 - Obtaining 60 % (or more) interesting references was:

easy ☐ difficult ☐ impossible ☐

NOTE : Starting from here, only one answer per SDI user is needed. Users having more than one profile need only to answer once (on the Questionnaire for the lowest profile code); they need not recopy their answers onto their subsequent Questionnaires.

Questionnaire No. 3 - page 3

9.1	Since you have been using computerized documentation (11 months now for most of you) has your overall problem concerning availability of journal articles: decreased ☐ ; stayed the same ☐ ; increased ☐ Comments as appropriate:	
9.2	Since you have been using computerized documentation (11 months now for most of you) has your overall problem concerning translation of articles: decreased ☐ ; stayed the same ☐ ; increased ☐ Comments as appropriate:	
10	What is the practical impact of a computerized documentation service on your information gathering behavior? - Did the computerized documentation service decrease the time needed for your documentation?	
10.1	Yes ☐ No ☐	
10.2	- If possible give time estimates, in hours/working week - before computerized documentation: _____ hours/week - with computerized documentation: _____ hours/week This question is important for the the evaluation. Please take the time to give the best estimates you can provide. (There is no obligation to use round figures if the quality of your estimates justifies the presence of decimals).	

Questionnaire No. 3 - page 4

10.3

Please mark one of the following possibilities **insofar as the areas covered by your profiles are concerned:**

1. You rely now solely on computerized documentation to locate suitable references ..... ☐
2. You feel that computerized documentation is an extremely profitable supplement to your manual documentation ..... ☐
3. You judge computerized documentation systems to be frequently useful but not extremely profitable. ☐
4. You judge such systems occasionally to be useful.. ☐
5. You find a computerized documentation system of no utility and you prefer to stick to traditional documentation methods ..... ☐

10.4

Stated in a more general way, do you feel better informed now than before the computerized documentation service?

Yes ☐

No ☐

10.5

What is the total number of researchers making use of the output of your **profiles** (completely or partially, including yourself)?

1 ☐

2 ☐

3 ☐

4 ☐

more (specify):

Questionnaire No. 3 - page 5

11

You have now been aware of computerized documentation and its advantages for 11 months*. Though you already gave a price estimate in Questionnaire No. 2, would you confirm (or modify) this by noting here how much you (or your organization) would be willing to pay to have the computerized documentation service continued (values are per year and per profile) ?

(* : for most of you)

numerical value	currency

Comments as appropriate:

12

Please mark one or several aspects in which improvements are needed in your opinion :

1 - The coverage of the database is insufficient ... ☐

2 - Pertinent references are not retrieved..... ☐

3 - Obtaining original documents or copies of the retrieved references is difficult..... ☐

Questionnaire No. 3 - page 6

13	Your total number of profiles for SDI Services is now: ...	
14	If you have more than one profile, please give the reason: 1. more than one field of research interest since the very beginning of the SDI service: ..... ☐ 2. you started with one profile, but new field(s) of research interest (how many, please specify:) appeared during the SDI experiment, and the corresponding profile(s) has (have) been added: ..... ☐ 3. too many references from one profile, obliging you to "split" it into two (or more:) more specific profiles: ..... ☐ 4. If necessary: comment points 1, 2, 3 or describe your particular case (please be as clear and complete as possible, but mention research fields in very general terms only).	

Questionnaire No. 3 - page 7

15

Please mark one or several aspects in which improvements are needed in your opinion (one answer per SDI user):

1 - Need for retrospective searches..... ☐

2 - The "Profile Design Manual" is difficult to use ☐

3 - The search logic is unsatisfactory..... ☐

4 - User training is not adequate..... ☐

5 - The format of the computer output is not satisfactory..... ☐

6 - The time needed for correcting the profiles is long..... ☐

7 - User's feedback system takes too much time..... ☐

8 - Not enough liaison officers to provide quick help to users..... ☐

9 - You have no criticism to make and you feel that the present system is satisfactory in your opinion ..... ☐

16

To be complete, a computerized documentation service should offer both of the following possibilities:

- documentation for current awareness (generally called "selective dissemination of information"(SDI))
- retrospective documentation (i.e. searches going back over several years for a specific subject).

For such retrospective documentation (especially useful when one has frequently changing research interests), how many years of retrospective searching would seem reasonably satisfactory in your opinion?

2 years	4 years	6 years	more (specify)
☐	☐	☐	-------- years

Questionnaire No. 3 - page 8

17 17.1	Please <u>comment freely</u> on any subjects related to a computerized documentation service and any possible improvements you would like. (Where appropriate, a clear separation of your different ideas will facilitate their subsequent handling.)

Questionnaire No. 3 - page 9

17.2 a.	- Should computerized documentation be continued in your case? Yes ☐ No ☐	
b.	- If you have more than 1 profile, give a specific (yes/no) answer for each of your profiles: code: answer: code: answer: code: answer: code: answer:	
c.	- Why and under what conditions (if not already explained in your general comments, see question 17.1)?	
18	Date: Signature:	

We hope that these questionnaires did not take too much of your time. Please carefully check to make sure that all questions have been answered. This is most important for proper evaluation and improvement of our documentation services. Your cooperation in greatly appreciated.

Thank you.

REFERENCES

1. Intergovernmental Conference on Scientific and Technological Information for Development. UNISIST II. Main Working Document. Unesco, Paris, May 1979 (PGI-79/CONF.201/COL.5).

2. Brown, J.E. "Developing SDI services; Canadian contribution to UNISIST". Canadian Library Journal, Vol. 32, December 1975, p. 451-455.

3. (a) Gietz, R.A. "Chemical selective dissemination of information service in Argentina"; (b) Raizada, A.S. and Satyananayana, R. "Indian experience in computer-based SDI service using CAN/SDI software package". Communications in the International Seminar on Selective Dissemination of Information and Technical Information, Ottawa, Canada, 2-4 October 1978; (c) Molino, E. "Evaluation report on the establishment of a CAN/SDI-based information service in Mexico". Prepared for Unesco under contract No. SC/RP 250.301. CONACYT, Mexico City, December 1976.

4. Canada Institute for Scientific and Technical Information. CAN/SDI Profile Design Manual. 5th edition. National Research Council, Ottawa, 1975. 44 p. plus annexes (CNRC No. 14676).

5. Schieber, W. ISIS: Integrated Scientific Information System. General Description of an Approach to Computerized Bibliographic Control. International Labour Office, Geneva, 1971. 115 p.

6. Tocatlian, J. "International information systems". Advances in Librarianship, Vol. 5, p. 1-60. Academic Press, New York, 1975.

7. One can find useful general considerations in this area in S. Nora and A. Minc. L'Informatisation de la Société. La Documentation Française, Paris, 1978.

8. CDS/ISIS: A General Description (Draft). Unesco, Paris, 1978.

9. UNISIST. Guidelines for the Conduct of User Studies (Revised Draft). Prepared by G. Ford, K. Wilson-Davis and J. Becker. Unesco, Paris, 1978. 158 p. plus annexes (PGI-78/WS/28).

10. UNISIST. Guidelines for the Evaluation of Information Systems and Services. Prepared by F.W. Lancaster. Unesco, Paris, 1978. 180 p. (PGI-78/WS/18).

11. Lemaire, F.C. Coûts et rentabilités de l'informatique. Dunod, Paris, 1972.

12. Research and Development Statistics 1973-1974. National Committee on Science and Technology, New Delhi, May 1975; Unesco Statistical Yearbook, 1975. Unesco, Paris (Section of statistics on science and technology).

13. Whitehall, T. Personal Current Awareness Service: A Handbook of Techniques for Manual SDI. The British Library, London, 1979 (R&D Report No. 5502).

BIBLIOGRAPHY

GENERAL REFERENCES ON INFORMATION RETRIEVAL SYSTEMS

Atherton, P. Handbook for information systems and services. Paris: Unesco, 1977.

Chaumier, J. Systèmes informatiques de documentation. Paris: Entreprise Moderne d'Edition, 1972.

Eyre, J. and Tonks, P. Computer and systems: an introduction for librarians. London: Clive Bingley, 1971.

Hall, J.L. On-line information retrieval sourcebook. London: ASLIB, 1977.

Lancaster, F.W. Information retrieval systems: characteristics, testing and evaluation (2nd edition). New York: Wiley, 1978.

Lancaster, F.W. and Fayen, E.G. Information retrieval on-line. Los Angeles: Melville Publishing Company, 1973.

Lindquist, M. The dynamics of information search services. Stockholm: Institut Royal de Technologie, 1978.

Meadow, Ch. T. The analysis of information systems (2nd edition). Los Angeles: Melville Publishing Company, 1973.

Salton, G. Automatic information organization and retrieval. New York: McGraw Hill, 1968.

Tcherny, A.I. Introduction à la théorie de la recherche documentaire. Moscow: VINITI, 1975.

Trystram, J.P. La documentation automatique. Paris: Dunod, 1971.

Van Slype, G. et al. Systèmes documentaires et ordinateurs. Paris: Les Editions d'Organisation, 1973.

Vickery, B.C. Information systems. London: Butterworths, 1973.

REFERENCE ON NETWORKS

Abramson, N. and Kuo, F.F. (ed.). Computer communication networks. Englewood Cliffs, N.J.: Prentice Hall, 1973.

GENERAL REFERENCES ON EVALUATION OF COMPUTERIZED INFORMATION SERVICES

Bassano, J.C. "Evaluation du coût et de l'efficacité des systèmes de documentation automatique". Automatisme (Paris), Vol. 19, No. 2, 1974. p. 90-95.

Cost of scientific and technical information and documentation systems. Luxembourg: Commission of the European Communities, Directorate General "Scientific and Technical Information and Information Management", 1976, No. FB975.

Farradane, J. "Evaluation of information retrieval systems". Journal of documentation, Vol. 30, No. 2, 1974. p. 195-209.

Grober, V.D. "O vybore kriteria ocenki dejatel'nosti informacionnyh služb (Choice of an evaluation criterion for the activities of information services). Naučno-tehn. inform. (Moscow), series 1, No. 1, 1970, p. 7-9.

King, D.W. and Bryant, E.C. The evaluation of information services and products. Washington D.C.: Information Resources Press, 1971.

Kopff, G. et al. "L'Evaluation des systèmes documentaires; problèmes théoriques et résultats attendus". Documentaliste, No. Spécial, 1974.

Krevitt, B.I. and Griffith, B.C. "Evaluation of information systems: a bibliography, 1967-1972". Information Part 2, Vol. 2, No. 6, 1973. p. 1-34.

UNISIST. Guidelines for the evaluation of information systems and services. Prepared by F.W. Lancaster. Paris: Unesco, 1978. No. PGI-78/WS/18.

Lancaster, F.W. "Evaluation and testing of information retrieval systems". In: Encyclopedia of library and information science, Vol. 8. New York: Marcel Dekker, 1972.

Samuelson, K. et al. Evaluation of information retrieval systems and functions. Stockholm: Nordfosk, 1974.

Solominin, V.M. "O pokazateljah effectivnosti dejatel'nosti informacionnyh organov" (Effectiveness parameters for activities of information centres). Naučno-tehn. inform. (Moscow), series 1, No. 5, 1974. p. 3-7.

Vickery, A. and Batten, A. Large-scale evaluation study of on-line and batch computer information services. University of London: Library Resources Co-ordinating Committee, 1978.

REFERENCES ON SELECTIVE DISSEMINATION OF INFORMATION

Bayard, T. and Persoz, J. "Comparison between CACON and CASIA files for development of new SDI service in 1977". J. Chem. Inf. Comp. Sci., Vol. 17, No. 2, 1977. p. 89-94.

Blase, N.G. "An SDILINE evaluation". Bull. Med. Libr. Ass., Vol. 64, No. 4, 1976. p. 412-415.

Bourne, Ch. P. "Improvements in the coupling of SDI system output with document delivery systems". J. Chem. Inf. Comp. Sci., Vol. 16, No. 1, 1976. p. 27-30.

Brown, J.E. "Developing SDI services; Canadian contribution to UNISIST". Canadian Library Journal, Vol. 32, December 1975. p. 451-455.

Butterly, E. "Improving SDI search profiles", Inf. Proc. Man., Vol. 11, No. 8/12, 1975. p. 189-200.

CONICYT. Servicio de Diseminación Selectiva de Información (DSI)/evaluación. Santiago (Chile): Comisión Nacional de Investigación Científica y Tecnológica, 1976 (Series Información y Documentación No. 14).

Cossmann, R.S. "Selective dissemination of information: a state of the art for the past five years". Libri, Vol. 27, No. 2, 1977. p. 125-135.

Dammers, H.F. "Economic evaluation of current awareness systems" In: EURIM: a European Conference on Research into Management of Information Services and Libraries, p. 107-124. London: Aslib, 1974. Also appeared in Cah. Docum., Vol. 27, No. 3/4, 1973. p. 37-61.

Davison, P.S. "Operating large-scale, broad-coverage CA and SDI services". Information science (London), Vol. 6, No. 1, 1972. p. 15-31.

Evans, L. Degree of user participation in SDI profile generation. London: Institution of Electrical Engineers, 1973. No. INSPEC-R-73-12.

Gluchowicz, Z. Selective dissemination of information and retrospective searches. Computer-based information services from RIT. Stockholm: Royal Institute of Technology (Library), 1973. No. TRITA-LIB-4023.

Gryaznov, N.I. and Saplina, T.B. "Osobennosti informacionnogo obespečenija NIORK na urovne nizovei informacionnoj služby NII; KB" (Features of information provision for research and development at the local level in scientific research institutes and design bureaux). Naučno-tehn. inform. (Moscow), series 1, No. 6, 1977. p. 1-5.

Harvey, S.A. A comparative study on the performance of two agricultural data bases in a computerized current awareness service. Wageningen (Netherlands): PUDOC, 1979.

Canada Institute for Scientific and Technical Information. CAN/SDI Profile Design Manual. 5th edition. Ottawa: National Research Council, 1975. No. CNRC-14676.

Kasiviswanadham, S. and Raizada, A.S. "Chem/SDI profiling". Ann. Libr. Sci. Doc. (Delhi), Vol. 23, No. 3, 1976. p. 223-234.

Leggate, P. et al. The BA Previews Project: development and evaluation of a mechanized SDI service for biologists. London: Office for Scientific and Technical Information (now integrated with the Research and Development Department, The British Library), 1973. No. OSTI - 5140.

Leggate, P. "Computer-based current awareness services". Journal of Documentation, Vol. 31, No. 2, 1975, p. 93-115 (Survey of literature in the field. Includes a bibliography of 152 entries).

Martin, J.K. and Parsons, R.G. "Evaluation of current awareness service for physics and astronomy literature". J. Am. Soc. Inf. Sci., Vol. 25, No. 3, 1974, p. 156-160.

Mekhtiev, D.M. et al. "Modelirovanie sistem izbiratel'nogo rasprostranenija informacii kak sredstvo ocenki effektivnosti" (SDI service modelling for efficiency evaluation). Naučno-tehn. inform. (Moscow), series 2, No. 2, 1974. p. 19-21.

Olive, G. et al. "Studies to compare retrieval using titles with that using index terms". Journal of Documentation, Vol. 29, No. 2, 1973. p. 169-191.

Persson, O. and Hoglund, L. Evaluation of a computer based current awareness service for Swedish social scientists. Stockholm: Royal Institute of Technology (Library), 1975. Research report, No. 29.

Presanis, A. and Brookes, B.C. COMPENDEX evaluation project: a study SDI services in engineering provided by COMPENDEX and of engineers' uses of and reactions to these services. London: Office for Scientific and Technical Information (now integrated with the Research and Development Department, The British Library), 1973, No. OSTI-5182.

Raizada, A.S. and Devadasson, F.J. "Chem/SDI service: performance of CAN/SDI software". Ann. Libr. Sci. Doc. (Delhi), Vol. 23, No. 2, 1976, p. 179-188.

Robredo, J. et al. The AGRIS data base as a support for the Selective Dissemination of Information Service BIP/AGRI. Brasilia: Empresa Brasileira de Assistência Técnica e Extensão Rural, Sistema Nacional de Informaçao e Documentaçao Agrícola, 1976, No. DOC/TECH/76/019 (Rev. 1).

Sprague, R.J. and Frendenreich, L.B. "Building better SDI profiles for users of large, multidisciplinary data bases". J. Am. Soc. Inf. Sci., Vol. 29, No. 6, 1978. p. 278-282.

Stern, L.W. et al. "Promotion of information services: an evaluation of alternative approaches". J. Am. Soc. Inf. Sci., Vol. 24, No. 3, 1973. p. 171-179.

Whitehall, T. Personal current awareness service: a handbook of techniques for manual SDI. London: The British Library, 1979. R&D Report No. 5502.

Wilson, J.P. and Wood, M.J. "An SDI experiment using COMPENDEX: an evaluation". New Zealand Libr., Vol. 41, No. 1, 1978. p. 4-10.

Yunis, S.S. "The implementation, evaluation and refinement of a manual SDI service". Bull. Med. Libr. Ass., Vol. 61, No. 1, 1973. p. 4-14.

REFERENCES ON STATISTICS

Hald, A. Statistical theory with engineering applications. New York: Wiley & Sons, 1967.

Parreins, G. Techniques statistiques. Paris: Dunod, 1974.